The Dissertation Committee for Matthew Leonard Cohen
certifies that this is the approved version of the following dissertation:

Bargaining and Fighting in the Moonlight

Committee:

R. Harrison Wagner, Supervisor

Tse–min Lin, Supervisor

Peter Trubowitz

Patrick McDonald

James Granato

Bargaining and Fighting in the Moonlight

by

Matthew Leonard Cohen, B.A.; M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2011

For Iris.

Acknowledgments

First, I want to thank the members of my dissertation committee. Peter Trubowitz encouraged me to make my arguments more forceful and to add to them more historical detail, and gave me positive encouragement in moments when I needed it most. Pat McDonald pushed me to think about the role of writing this dissertation in my professionalization as an academic, helping me to better articulate the relationship between my work and the broader literature to which it belongs. Jim Granato helped me to think more carefully about how to link my empirical and theoretical chapters.

My deepest gratitude is reserved for my supervisors, Tse–min Lin and R. Harrison Wagner. Tse–min helped a very philosophically inclined student to see that there are fascinating philosophical questions about the process of observing and measuring empirical reality, and great pleasures to be gained in doing so. Harrison challenged me to be a more analytically rigorous thinker, and inspired me to be sometimes more humble and other times more bold in my theorizing. It would be difficult to overstate their positive influence on my intellectual development.

I also want to thank my undergraduate advisor, Peter Cocks, who in his seminars at Simon's Rock College first kindled in me an interest in the study of politics. While Peter was skeptical about the increasing dominance

of statistical and game theoretical modeling in political science, a trend that this dissertation is certainly a part of, I hope he will be able to recognize his own influence under all of the mathematical equations.

In the early stages of writing, I was lucky to have the support and feedback of Eunjung Choi, Oddysseas Christou, Scott Garrison, Feng-yu Lee, Ayesha Ray, Laura Seay, and Jongseok Woo. I am grateful to them for all of their help.

My students at Bard High School Early College in Queens have asked consistently penetrating and insightful questions about political science as a discipline and my own contributions to it. I owe much of the clarity of my writing — that is, where there is such clarity — to the conversations I had with them inside and outside of the classroom.

My parents, Mary Lind–Cohen and Robert Cohen, are responsible for my earliest intellectual development, as well as my earliest introductions to this thing called politics. They pushed me to leave home and go to college at a time when I was far too provincial and complacent in my thinking. Without these early interventions, I would never have dreamed of pursuing a PhD in any field.

Finally, I would not have been able to finish this project without the love, support, patience, and gentle prodding from my wife, Iris. I cannot begin to express how happy I am to have found you, Iris, and how lucky I feel to have you in my life.

Bargaining and Fighting in the Moonlight

Publication No. _____

Matthew Leonard Cohen, Ph.D.
The University of Texas at Austin, 2011

Supervisors: R. Harrison Wagner
Tse–min Lin

"Audience costs" models of international relations suggest a purely informational role for domestic politics in conflict settings. Here, domestic politics serve as a rich signal of belligerents' true intentions, allowing them to more quickly resolve disagreements, decreasing the likelihood and duration of war. But if belligerents can have different *beliefs* about *publicly available* information, then domestic politics might confuse rather than clarify conflict situations, increasing the likelihood and duration of war. I present empirical evidence of conventional "audience costs" models' shortcomings in explaining the dynamics of the US counterinsurgency efforts in Iraq and the response of Iraqi insurgents to those efforts. I then develop a formal model to show how differences in beliefs between insurgents and counterinsurgents about domestic political audiences in Iraq may have contributed to the prolonged nature of the conflict. I argue that the underlying cause of the conflict's duration is disagreement between belligerents about whether and how Iraqi civilians

contribute to a successful counterinsurgency, leading belligerents to disagree not only before fighting about who is likely to win, but during fighting about who is actually winning.

Table of Contents

Acknowledgments **v**

Abstract **vii**

List of Tables **xii**

List of Figures **xiii**

Chapter 1. Introduction **1**

1.1 Agreeing to disagree about the costs of war 1
1.2 Why should we care what they think? 12
1.3 The second Gulf War 24
1.4 The first Gulf War 28
1.5 Structural outline 32

Chapter 2. Literature review **37**

2.1 Moving beyond the Realist view of domestic politics 38
2.2 Domestic audience costs 43
2.3 Bargaining while fighting: Mutual optimism and its remedy . . 52
2.4 Explaining protracted wars: Domestic audiences, mutual optimism, and credible commitments 58

Chapter 3. Research design **62**

3.1 The empirical implications of theoretical models 62
3.2 The theoretical implications of empirical models 64
3.3 Data 66
3.3.1 General requirements of the data 67
3.3.2 Public opinion 68
3.3.3 Hostility 70

3.4 Empirical Model 80
3.4.1 Estimation 84
3.4.2 Interpretation 85
3.5 Theoretical implications 87

Chapter 4. Were Bush's critics a fifth column? 89
4.1 An emboldening effect? 89
4.2 Signaling resolve 94
4.2.1 Bystanders and followers, accountability and credibility 101
4.2.2 Can an uninformed public inform? 103
4.3 Analysis 104
4.4 Conclusion 114

Chapter 5. A Random Walk in the Moonlight: Divergence and Convergence of Beliefs in Insurgency and Counterinsurgency 119
5.1 Motivation 120
5.1.1 Two views of insurgency and counterinsurgency 123
5.1.1.1 The "population centric" approach 123
5.1.1.2 Enemy–centric 126
5.1.2 Antecedents 128
5.1.3 Practical implications of disagreement 130
5.1.4 Rationalist theories of war 130
5.1.5 Overestimating the probability of future success 132
5.1.6 The belligerent's ruin 133
5.2 Divergent beliefs in insurgency and counterinsurgency 139
5.2.1 Counterinsurgents' beliefs about choosing C 144
5.2.2 Insurgents' beliefs about counterinsurgents choosing C . 146
5.2.3 Counterinsurgents beliefs about choosing G 147
5.2.4 Insurgents' beliefs about counterinsurgents choosing G . 149
5.2.5 How settlements are reached 151
5.3 A model of bargaining, fighting and learning during the US led counterinsurgency in Iraq 161
5.4 Conclusion 168

Chapter 6. Conclusion **172**
6.1 Summary of findings 172
6.1.1 No emboldenment effect in Iraq 172
6.1.2 Disagreements about success and failure 174
6.2 Implications 174
6.3 Future research 176

Bibliography **179**

Vita **192**

List of Tables

4.1 Results of Granger test 108

5.1 Equally certain about means and about strategy 166
5.2 Uncertain about means and certain about strategy 166
5.3 Certain about means and uncertain about strategy 167

List of Figures

3.1 Regular intervals, constant approximation 71
3.2 Public Opinion 72
3.3 Total level of aggression or cooperation, US to Iraq 81
3.4 Total level of aggression or cooperation, Iraq to US 82

4.1 Decomposition of Forecast Error Variance for US Hostility . . 108
4.2 Decomposition of Forecast Error Variance for Iraq Hostility . . 108
4.3 Decomposition of Forecast Error Variance for Public Opinion . 109
4.4 Impulse Response Functions from Public Opinion 109
4.5 Impulse Response Functions from US Hostility 110
4.6 Impulse Response Functions from Iraq Hostility 110

Chapter 1

Introduction

1.1 Agreeing to disagree about the costs of war

President George W. Bush's televised speech on January 10th, 2007, announced a significant change in America's approach to its counterinsurgency efforts in Iraq (Sanger, 2007). Because it involved the deployment of an additional 20,000 troops to Iraq and the extension of the tours of duty of the majority of American troops already stationed in the country, the plan he described had already been dubbed by the media as "the surge" well in advance of Bush's speech (Sanger and Gordon, 2006). But more important than its signaling America's increased commitment to the war in Iraq, the speech indicated a radical change in thinking about the nature of the war itself — one so fundamental that concepts such as "success" and "failure" were ultimately redefined (Kilcullen, 2009). It is my contention that explaining this change requires a novel addition to "rationalist" explanations of protracted conflicts such as the war in Iraq: Belligerents may disagree not only *before fighting* about who is likely to win a war, but also *while fighting* about who is, in fact, winning. Where the latter kinds of disagreement arise, conflict is more protracted than it otherwise would be. To allow for such disagreements requires a move away from the widespread convention in the rationalist literature of

treating all disagreements as essentially *informational* in nature.

Much of the literature on "bargaining while fighting" (Wittman, 1979; Wagner, 2000; Powell, 2004; Smith and Stam, 2004) conceptualizes war as a result of, and a means for belligerents to overcome, *mutual optimism*. Where mutual optimism is a problem, some parties to a conflict may initially overestimate their own chances of winning a war against the others, thus causing them to think they can demand more from a negotiated settlement than their adversaries believe they should have to concede. To overcome mutual optimism, belligerents fight, directly testing the validity of their enemies' and their own threats. War is thus not only the result of disagreement, but ultimately its solution: As belligerents fight, they learn more reliable information about their own and each others' actual likelihoods of winning. As they learn, and as their expectations and beliefs about winning converge, belligerents can ultimately reach a mutually agreeable settlement.[1]

Such learning can be understood to occur as a result of belligerents inflicting, suffering, and/or avoiding the costs associated with war. As belligerents suffer greater than anticipated costs, and/or impose on their adversaries lesser than anticipated costs, they revise downward their expected likelihoods of winning the war.[2] Ideally, the process of suffering and imposing

[1] See chapters 2 and 5 for more thorough discussions of the logic behind arguments based on the concept of mutual optimism.

[2] And, of course, the opposite should also be true. That is, as belligerents suffer lesser than anticipated costs, and/or impose on their adversaries greater than anticipated costs, they revise upward their expected likelihoods of winning the war.

costs through battle should ultimately result in bringing each belligerent's expectations in line with the others'. But this process of convergence requires that belligerents agree on what costs are important, and on who has suffered the most. As I will suggest below, the structure of these costs can be quite complicated, and thus arriving at a mutually agreeable measure of them is often quite difficult.

Broadly speaking, belligerents seek to impose on one another and avoid for themselves two kinds of costs: The first consists of direct, material costs, such as the destruction of property, the seizure of territory, and civilian and military casualties. The second consists of indirect, political costs, such as the loss by political elites of power and prestige. The two are not mutually exclusive — much of the political costs of war flow from material costs, as political elites are punished by their constituents for their failure to avoid them. There is, in fact, an argument to be made that in many modern wars it is only insofar as they translate into political costs that material costs can explain war outcomes at all. This is because political elites whose decisions determine conflict outcomes are usually well insulated from any direct experience of the material costs of war. It is thus only when the political masses who do directly experience such costs can hold their leaders accountable for their decisions that these costs have any bearing on the nature of their leaders' decisions.[3] As I argue below, which of these costs matter, as well as when and how they matter,

[3]This insight is behind the use of principle–agent models from economics to explain war outcomes, such as (Downs and Rocke, 1994), discussed in greater detail in chapter 2.

are questions of great importance to academics and policymakers alike, and that there are no easy, definitive answers to these questions can help explain the dynamics of some protracted conflicts. Where it is possible to disagree over fundamental questions such as how to measure costs and success, it is more difficult for belligerents to arrive at war terminating settlements.

Most research on the causes and consequences of mutual optimism has focused on the role of informational asymmetries in disagreements leading to war (Fearon, 1995). In models of conflict based on the existence of informational asymmetries, belligerents have private information about their own capabilities and incentives to misrepresent to each other what those capabilities are and the extent to which they are willing to use them. Mutual optimism, then, is a result of belligerents' inability and/or unwillingness to accurately and convincingly communicate to each other their capacity for imposing and suffering the costs associated with war. But modeling the relationship between the costs of war and mutual optimism as purely *informational* is problematic in that it suggests that overcoming mutual optimism is easier than it often is. In such models, because the costs of war are essentially *public information*, equally accessible by all belligerents and, more importantly, equally *significant* to all of them, it is implied that belligerents' expectations about the future converge very quickly once war commences. Such models therefore have little to say about the duration of war in general, and protracted conflict in particular. Fey and Ramsay (2007) have gone so far as to suggest that the process of overcoming mutual optimism is, in fact, so quick and easy that belligerents

should always be able to overcome informational asymmetries leading to war through standard bargaining, without ever having to fight or pay any costs whatsoever (though they seem to believe this to be an indictment of the concept of mutual optimism in general, rather than of the convention of modeling mutual optimism as purely informational).

These problems stem from the widespread practice in game theoretic modeling of assuming that actors have "common priors." In the Bayesian framework, what an actor knows is due to a combination of her prior beliefs and her observations of new facts. This partitioning implies that when two actors disagree about something, it can be either because they began with different prior beliefs or because they have observed different new facts. Since Harsanyi's introduction of the concept of the Bayesian equilibrium (Harsanyi, 1967, 1968a,b), the convention is to assume that the latter is the case, and this assumption has developed into a kind of dogma among many formal theorists in the social sciences (Fey and Ramsay, 2006). But while it is often a useful assumption, it is far from a necessary one. Despite arguments to the contrary, it is possible for two actors with access to the same information to disagree about the meaning and significance of that information. In other words, it is often the case that people to "agree to disagree," and formal models should be able to accommodate such disagreement (Morris, 1995). A failure to do so can result in arguments that are too limited in their scope or, as in Fey and Ramsay (2007), nonsensical in their implications.

Smith and Stam (2004) have thus suggested an alternative (or addi-

tional) formal model of mutual optimism, allowing for the possibility that even when belligerents have access to the same information, they might initially interpret that information quite differently. If, for example, belligerents disagree about whether air or naval power is the key to ultimate success in the war they are preparing to fight, then even if they have access to the same information about their relative capabilities and resolve — that is, even if they are both completely honest with each other about their ability and willingness to fight — the belligerent with the larger navy and the belligerent with the larger air force might both believe they are more likely than the other to win. War then becomes a way not (or not only) of communicating otherwise private information, but of determining whose interpretation of the initial, publicly available information was correct. Belligerents that find themselves performing more poorly in battle than they first expected might thus ultimately be compelled to take their adversaries' original interpretation of their relative strengths and weaknesses more seriously.

Given that belligerents may "agree to disagree," one might reasonably ask about what, exactly, they can disagree. As I argue below, the change in America's approach to counterinsurgency in Iraq (and, earlier, in Vietnam), suggests an answer to this question that is not . *Belligerents may disagree not only in their interpretations of information available prior to war, but also in their interpretations of information that war reveals.* Briefly,[4] America's changes in approach to counterinsurgency are based not only on different tac-

[4]I will explore this argument in greater depth in chapter 5.

tics, but also on different sets of *metrics for success*, and there is still broad disagreement among American pundits and policymakers about which of these metrics are better. If American counterinsurgents can disagree among themselves about what constitutes success and failure, then surely Iraqi insurgents can similarly disagree with American counterinsurgents as well. When observers can disagree on how success and failure should be measured, this poses a problem for many explanations of war termination that treat war as a means of overcoming mutual optimism. Given different metrics — i.e. different understandings of the costs associated with war — belligerents might look at the same battle and legitimately disagree about whether its outcome contributed to or subtracted from their own ultimate success in the larger war; in other words, they might disagree about *who had actually won the battle.* Belligerents could then learn the wrong lessons from fighting, increasing rather than decreasing mutual optimism, and therefore prolonging rather than shortening the conflict.

In what follows, I argue that such disagreements are common among counterinsurgents and insurgents, and that these disagreements contribute to the protracted nature of many such conflicts. An insurgency's or counterinsurgency's success may depend in some part upon the efforts of insurgents and counterinsurgents to attract and keep supporters and recruits from the civilian population, and the behavior of insurgents and counterinsurgents can, intentionally or unintentionally, either aid or undermine such efforts. But insurgents and counterinsurgents cannot, based upon the outcome of any battle

viewed in isolation, determine how it will affect their prospects for ultimate success. They first of all will not be able to gauge the success of their respective efforts to win over the civilian population, as their success in this regard will depend largely on factors that neither side can directly observe or definitively predict — in particular, shifts in the civilian population's opinions about the insurgents, the counterinsurgents, and the conflict in general. Thus, whether any battle has aided or undermined one side's or the other's efforts to win over the civilian population is in the short-run a matter of subjective interpretation guesswork, and therefore a potential cause for disagreement. But they may also disagree about whether the civilian population is important to their objectives at all, and again the outcome of any given battle is unlikely to provide much in the way of evidence for one perspective or the other. Under some circumstances these kinds of disagreement can undermine belligerents' ability to learn from war accurate information about each others' capabilities and resolve, leading to divergence rather than convergence of their beliefs and expectations about winning. Where such disagreement is possible, it takes much longer for belligerents' expectations to come into alignment than standard models of mutual optimism suggest.

America's involvement in the war in Vietnam could be seen as a case in point. In an interview after the end of the conflict, the Vietnam People's Army's Commanding General Vo Nguyen Giap explained that in his opinion "there is no such thing as a single strategy. Ours is always a synthesis, simultaneously military, political, and diplomatic — which is why, quite clearly, the

Tet offensive had multiple objectives," and at the time of the Tet Offensive, Giap claimed in a radio address that one of these objectives was to "make the American people more aware of the errors and setbacks of the Johnson administration in the aggressive war in Vietnam" (Willbanks, 2007). In other words, the Vietnam war was, as US President Lyndon Johnson famously claimed, a fight for "hearts and minds," but in Giap's opinion it was the hearts and minds of Americans, not Vietnamese, that mattered. Johnson was not at all ignorant of Giap's opinion. In his "San Antonio Formula" speech of September, 1967 (about four months before Tet), Johnson conceded

> They still hope that the people of the United States will not see this struggle through to the very end. As one Western diplomat reported to me only this week — he had just been in Hanoi — "They believe their staying power is greater than ours and that they can't lose." A visitor from a Communist capital had this to say: "They expect the war to be long, and that the Americans in the end will be defeated by a breakdown in morale, fatigue, and psychological factors." The Premier of North Vietnam said as far back as 1962, "Americans do not like long, inconclusive war... Thus we are sure to win in the end." Are the North Vietnamese right about us? I think not. No. I think they are wrong... (Willbanks, 2007)

This would appear to be an instance of the Americans and the North Vietnamese "agreeing to disagree" — surely it would be absurd to argue that Giap

had information about American public opinion that Johnson did not. By this interpretation, the Tet Offensive was a victory for the North Vietnamese insofar as it convinced Johnson to accept their definition of success and failure in the war — one that included an awareness of the pivotal role played by civilian populations, even those living half a world away from the war itself.

Theoretically, the present study is a contribution to and synthesis of ideas from the rationalist literatures on bargaining and learning while fighting, mutual optimism, and the interdependence of domestic and international politics. In the next chapter I will review these literatures and explain their relevance to the study of insurgency and counterinsurgency. But in the remainder of this introduction, I will make the case for a broader application of the theoretical framework I present and develop herein. Given what Clauswitz calls the "fog" or "moonlight" of war, i.e. "the general unreliability of information" that "often tends to make things seem grotesque and larger than they really are" (von Clausewitz, 1993), any conflict setting may provide belligerents ample opportunities not only for learning, but for learning things that are wrong or misunderstanding what they have learned, and what follows provides one possible framework for thinking about such misunderstandings and their implications.

By focusing on war's "moonlight," this study identifies a role for domestic political actors in international conflict that is significantly different from the one usually ascribed to them in the conflict bargaining literature. This difference is due in part to the above mentioned move away from the treat-

ment of disagreement as merely *informational*. The "audience costs" approach advocated by Thomas Schelling, James Fearon, and Kenneth Schultz, as well as the "selectorate" models developed by Beuno de Mesquita, et. al., treat domestic political actors as a means for revealing information that international political actors otherwise have an incentive to hide or misrepresent and for confirming information that international political actors otherwise have good reason to doubt (Schelling, 1960; Fearon, 1995; Schultz, 2001a; Bueno de Mesquita et al., 2003). By so doing, domestic political actors potentially allow international political actors to limit fighting or even avoid it entirely. In other words, the primary contribution of domestic political actors to international conflict settings, according to much of the conflict bargaining literature, is in providing greater informational clarity and thereby reducing conflict.

But by treating the kinds of disagreement that lead to conflict as originating not in informational asymmetries, but in different beliefs about how best to interpret the available information, a new role for domestic politics opens up. Here the contribution of domestic political actors to a conflict bargaining setting is potentially one of confusion rather than clarification, thus increasing the likelihood and/or duration of conflict. While I focus on how such confusion may occur in the course of an insurgency and counterinsurgency, it is my contention that it is likely in other conflict settings as well.

To make this case, I begin with a brief discussion of the rationalist approach to thinking about politics and its usefulness in bringing clarity to many questions about conflict in general, and counterinsurgency/counterterrorism in

particular — a clarity that is sadly too often missing from public discourse on American foreign policy.[5] I then identify some persistent puzzles, first in the study of insurgency and counterinsurgency, and second in the study of conflict in general, that the present study helps to explain. Finally, I outline the structure, methodology, and theoretical framework of the book.

1.2 Why should we care what they think?

The perceived disconnect between informed, non-academic discourse on foreign policy and contemporary IR scholarship has been the subject of a great deal of handwringing among academics and policymakers alike (Nye, 2009; Walt, 2005). Many have pointed to what they believe to be the obscurantism of IR scholarship as the primary cause of this disconnect, and to the increasing use of rational choice theory as one of the primary causes of said obscurantism. Stephen Walt, one of the louder of these voices, has bemoaned what he calls a growing "cult of irrelevance" in IR, scholars that he claims use increasingly complicated and arcane modes of mathematical reasoning to either 1) answer questions that are of no general interest or 2) merely rehash arguments that Walt believes he and his colleagues definitively settled decades ago. The result, Walt claims, is a body of scholarship that is of interest to no one who is

[5]The blanket term "rationalist" is, perhaps, ill-chosen, and its use quite possibly responsible for a number of persistent misunderstandings regarding the precise nature of arguments belonging to this subset of the IR literature Wagner (2010). But given that, as yet, no other satisfactory term has yet emerged to differentiate this literature from its alternatives, I will, for lack of a better option, continue to use "rationalist" to describe it.

concerned with real world issues.

> [R]ecent formal work has had relatively little to say about important real-world security issues. Although formal techniques produce precise, logically consistent arguments, they often rest on unrealistic assumptions and the results are rarely translated into clear and accessible conclusions. And because many formal conjectures are often untested, policymakers and concerned citizens have no way of knowing if the arguments are valid (Walt, 1999).

The implication is that IR scholars can regain the attention of foreign policy makers and pundits by abandoning the specialized, formal language of deductive logic and mathematics and returning to making their arguments in plain English, thus ensuring that their arguments will be both relevant and easily understood.[6]

This framing of the problem is only partially correct. It is certainly the case that rationalist arguments in IR — particularly those employing sophisticated game theoretical models — can be quite difficult for the non-expert to understand. In fact, some can tax the mathematical abilities of even those of us who are usually quite comfortable with the genre. But this difficulty does

[6]Of course, IR scholars often engage in scholarship of limited general interest and import *without* the use of sophisticated mathematical models, as evinced by the long and fruitless "absolute versus relative gains" debate of the 1990s between Realists and Neoliberal Institutionalists. And as it turns out, and despite Walt's beliefs to the contrary, these same models can be quite useful in definitively *deciding* such debates, as Robert Powell's seminal (and widely misunderstood) work on absolute and relative gains has done (Powell, 1991).

not prove their irrelevance. Rather it only suggests that when they *are* relevant we will have to take greater care to explain them to laypersons. In fact, I would contend that the absence from public debate of the "rationalist" perspective — with or without sophisticated mathematical models — diminishes discourse. To illustrate how this is the case, I will in this section examine a recent public debate concerning the intersection of domestic and international politics, and explain how the rationalist perspective could have had a positive influence on it.

One prominent feature of American political discourse in the summer of 2010 was the debate over the building of the Cordoba House (now Park51), a community center that Kuwaiti born Sufi Imam Feisal Abdul Rauf planned to build in downtown Manhattan.[7] The controversy focused on the center's location, a few blocks away from "Ground Zero" of the terrorist attacks on September 11th, 2001, and the proposed inclusion, among its facilities, of what was described by the organization alternately as a "mosque" and a "prayer space." Many critics of the project suggested that it was unwise to build so near the site of this tragic act of mass murder a place of worship dedicated to the very religion that had purportedly inspired its perpetrators.

Among the concerns central to this controversy — voiced by proponents and opponents of the Cordoba Initiative alike — was what message the center would send to the rest of the world, and in particular to those who might

[7]As of this writing, the project is still underway, but has not yet broken ground.

have sympathized with or even supported the September 11th attacks. Some suggested that the building of what they called the "Ground Zero Mosque" was symbolic of a defeat of American values, a monument to very the people who sought America's destruction. Conversely, others suggested that it was a celebration of those very same values, a testament to America's dedication to freedom and diversity, and an implicit repudiation of the very notion that American and Muslim values are fundamentally in conflict. On August 25th, comedian Jon Stewart lampooned all sides of this debate on his comedy news program "The Daily Show" (Stewart, 2010):

> See, one side says our weakness emboldens jihadis. The other side says our strength embitters jihadis. How about we try a new system, where we don't give a [expletive] about what they think? Why can't we do that? Let me give you an example. It's like wondering what your ex-girlfriend would think about the drapes you and your wife picked out for your apartment. No matter what pattern you pick, she's not going to like the drapes, because she's [expletive] crazy.

Stewart's remarks give expression to a number of related perspectives that appear to be shared by many, among not only the general public but also the political and media elite, regarding the nature of the wars in Afghanistan and Iraq and, broadly speaking the "war on terror": The conflation of explana-

tion and exoneration,[8] the assumption that our enemies are beyond reason,[9], and a variant of the pathetic fallacy where the attributes of individuals are ascribed to groups.[10] If we take seriously research suggesting that public opinion on issues of foreign policy is largely determined (or at least very heavily influenced) by the substantive focus, statements, and rhetoric of political and media elites,[11] then comments like Stewart's may be somewhat more significant than they at first appear. Our foreign policy is ultimately determined by public opinion, insofar as any foreign policy agenda can only be pursued so long as its proponents can get elected and reelected.[12] If public opinion of foreign policy is based on fundamental misunderstandings of the issues at stake, then foreign policy may very well suffer as a result. A brief (and, admittedly, humorless) discussion of the substance and implications of Stewart's joke may therefore prove instructive.[13]

What, exactly, is Stewart's argument? First, Stewart seems to be suggesting that the thoughts and motives of what he calls "jihadis" are so dif-

[8]Catalogued and explored with uncharacteristic clarity by Judith Butler (2002) in the aftermath of September 11th.

[9]More recently applied to the leadership of Iran, for example in Jason Lee Steorts's cover story for the October 23, 2006 issue of National Review, "Can Iran be Deterred?" (Steort, 2006, Steorts's answer: No.).

[10]A habit that is, unfortunately, just as common among political scientists as it is among laypersons, as evinced by the vast majority of the "Realist" literature in IR and its emphasis on the "unitary actor" model of the state.

[11]See Baum and Groeling (2009) for a comprehensive overview and impressive synthesis of this literature.

[12]See below and chapter 4 for further discussion of this connection.

[13]Note that while I single out Stewart in the analysis that follows, I do not mean to imply that the interlocutors he parodied were any more nuanced in their treatment of the issues at stake.

ferent from our own that they cannot be understood. They are crazy like his ex-girlfriend, and there is no sense in looking for reasons or reasoning behind their ideas and actions, because no such reason or reasoning exists. Second, regardless of whether it is possible to reach such an understanding, he appears to believe that there is little or no value — and in fact there may even be harm — in trying to do so. They are, again, crazy like his ex-girlfriend, thus their reasons and reasoning should not be taken seriously, because to do so would be to allow them a level of influence over our thinking that they do not deserve. Finally, he assumes a sufficient level of cohesiveness within the group he calls "jihadis," such that this group can be reasonably expected to possess attributes that we usually reserve for describing individuals — e.g. the very craziness that Stewart thinks makes understanding its thoughts and motives both impossible and unnecessary. Thus, a *group* that is made up of crazy individuals might very well be expected to think and behave *collectively* in a manner similar to that of a crazy *individual* in *isolation*.

How apt is this "crazy ex-girlfriend" metaphor, and how valid are the sentiments behind it? As it turns out, much (indeed, most) of contemporary scholarship in IR — especially, but certainly not exclusively, the literature commonly referred to as "rational choice" or "rationalist" — begins from a set of assumptions almost exactly opposite Stewart's. And while it is not necessarily a given that the starting assumptions of IR scholars are better than the alternatives proposed by non-IR scholars, there are good reasons to take them seriously, and to at the very least be skeptical of perspectives that

evince no awareness of them. That these assumptions, and their implications, play such an insignificant role in public discourse on foreign policy necessarily diminishes that discourse.

Few IR scholars would deny, of course, that the beliefs, motives, goals, and preferences of political actors may be quite different, such that one actor may well view the beliefs, motives, goals, and/or preferences of another as "crazy." But given those beliefs, motives, goals, and preferences, many IR scholars assume that there is some consistency, even predictability, in political behavior: Political actors are, at the very least, *purposive*, i.e. "they pursue their goals as best they can" (Lake and Powell, 1999; Elster, 2007; Wagner, 2010). Scholars working explicitly in the "rational choice" tradition are somewhat more specific about the relationship between reasons and actions, assuming that political actors will choose (or will act as if they have chosen) those actions that they think will maximize their own satisfaction (i.e. "utility") in the outcome. In either case, the basic idea is that even if political actors' goals make no sense to us, we can usually expect them to act in a manner consistent with achieving them. Even when political actors seem to violate these assumptions — especially the more strict assumptions of rational choice — experimental and observational evidence suggests that they deviate in ways that are systematic and predictable (Kahneman and Tversky, 1979). Thus, even when people are "crazy," as long as they are essentially *goal oriented*, we should usually be able to make some sense of the reasoning behind their actions; or, given an understanding of their reasoning, we should be able

to make some sense of their actions.

There is, then, good reason to want to understand political actors' reasoning. This is particularly true if we want to engage in political action ourselves. Many, perhaps most, of the settings for meaningful political action are essentially *strategic*, which is to say that outcomes depend on choices made by multiple actors with different and often contradictory goals, and that their choices are interdependent. Actors who want to achieve their most favored outcomes must try to anticipate the behavior of all of the other actors who might affect those outcomes, and condition their own behavior accordingly. But they must do so while keeping in mind the possibility that those other actors are anticipating their behavior, too, and conditioning their behavior accordingly. Even if we do not share others' reasons for action, knowing what their reasons are is exceedingly useful, as it allows us to formulate an appropriate strategy in anticipation and/or response.

That terrorists' and insurgents' reasons for *being* terrorists and insurgents are, from our perspective, crazy need not imply that their reasons cannot be understood, or that understanding those reasons would not be useful. In fact, that insurgents and terrorists are goal-oriented is not an assertion of great controversy among those who study insurgency and terrorism, even if there is some debate over what those goals may be. It is only in public discourse that discussion of these goals, or even acknowledgement that such actors have goals, is considered to be out of bounds. Even in the non-academic foreign policy community there are few (if any) approaches to thinking about counterinsur-

gency or counterterrorism that do not at some point presume that insurgents and terrorists are, like nearly all other political actors, essentially purposive in their behavior, and that there is utility in gaining some understanding of their purposes and the means available to them for pursuing them — so as to anticipate, respond appropriately to, or even preempt their chosen paths of action. The real controversies are over not whether terrorists and insurgents have reasons for their actions, or whether we should try to understand those reasons, but rather on what those reasons are, and what they imply about our own best courses of action.

Many of these controversies center on the question of whether and how insurgent and terrorist organizations are affected by problems of collective action and/or coordination. In collective action problems (Olson, 1971), cooperation for the sake of achieving an outcome that is both collectively and individually beneficial is made difficult or even impossible by a combination of individual selfishness and mistrust. Where there are costs associated with cooperation, and where full cooperation by all affected parties is not necessary in order to achieve the collectively beneficial outcome, then some might prefer to "free ride" on the efforts of others — in hopes of enjoying the beneficial outcome without paying the costs of cooperating to make it happen. If there are too many free-riders, and/or people who abstain because they expect free-riding from others, then the beneficial outcome will not be achieved. Thus the existence of a group of people actively working together towards some common goal presupposes that the group's members have overcome the collective

action problem — usually by either reducing the costs associated with participation, directly rewarding those who choose to participate, or imposing costs on non–participants.

In coordination problems (Schelling, 1960), cooperation is made difficult or even impossible by actors' inability to effectively and/or convincingly communicate to one another their own intentions. Assuming again that there are costs associated with cooperative behavior, anyone who wants to engage in such behavior might want assurance that a sufficient number of others would do the same, such that the preferred outcome would actually be achieved. The worst possible option would be to choose to engage in cooperative action when no one else does, and the desire to avoid such an outcome might be enough to convince everyone not to act at all. Thus, the existence of a group of people actively working together towards some common goal presupposes that the group's members have successfully overcome the coordination problem.

Modern, sovereign nation states are typically presupposed not to face collective action and coordination problems — that is, not internally. The provision of territorial security is certainly something that *could* be vulnerable to problems of collective action and coordination, but theorists since Thomas Hobbes have tended to assume that states exist to solve such problems, and that they are typically successful in doing so. Tax revenue provides states the capacity to offer selective incentives (salaries, benefits, etc.) for voluntary enlistment in their security services, and as a last resort they can make service mandatory. There are therefore fewer opportunities for free riding — shirking

of duty can be discouraged easily by threatening the withholding of salaries and benefits or the direct imposition of punishments such as incarceration — making collective action easier to ensure. Thus, for example, despite the risk of death associated with fighting a war, a state can get its soldiers to participate by providing counterbalancing rewards or threatening punishments more certain than death. Coordination problems are even less likely within organizations governed by sovereign nation states, as hierarchies with well-established chains of command and clear channels of communication minimize their occurrence. A soldier whose platoon has been ordered to advance can therefore usually be fairly certain that he will not be the only one who shows up at the front.

But non-state actors such as insurgent organizations are a different story entirely. First, only a subset of a given population is likely to be so profoundly unhappy with how, and perhaps more importantly by whom, it is ruled that it would be willing to participate in an insurgency. But second, and more importantly, it is not a given that even this subset will be able to transform itself into an insurgent organization. The violent removal and replacement of the current regime might be an outcome individually and collectively beneficial to the people who oppose it, but the fact of their wanting it to happen does not make it so. At a minimum the potential insurgents must solve the attending problems of collective action and coordination: Insurgencies can be quite dangerous, and the costs of participation quite high. Anyone who might prefer the overthrow of the government to the status quo might therefore also

prefer that someone else do the overthrowing. And even someone who is willing to shoulder the costs and risks of revolting might nonetheless want some assurance that when she arrives, armed, at the gates of the palace, she will not be arriving alone.

An insurgent or terrorist attack may be evidence, therefore, not only of decisions by an organizations' elites to conduct such an attack, but also of those organizations' members to join (and remain in) these organizations in the first place. It may, furthermore, be evidence of the decisions of civilians to either provide direct support to the insurgent organization or, at the very least, to withhold information about it from counterinsurgents. Lastly, it could be evidence of the failure of the counterinsurgent or counterterrorist organization to discourage fence-sitters — what David Kilcullen calls "accidental guerillas" (Kilcullen, 2009) — from joining or otherwise supporting the insurgents. In other words, counterinsurgents should be sensitive to the extent to which insurgents are sensitive to costs associated with collective action and coordination.

Thus, there are in fact very good reasons to "give a [expletive] about what they think." In the following sections I will explore in greater detail why it might be useful to do so, in the context not only of insurgency and counterinsurgency, but also that of conventional interstate conflict.

1.3 The second Gulf War

Among those in the US foreign policy community who advocate in favor of America's counterinsurgency efforts in Iraq and Afghanistan, a significant divide exists between those favoring a "population–centric" approach and those supporting more conventional, "enemy–centric" military strategies and tactics:[14] Should strategy focus on finding and killing insurgents, or on winning over the "hearts and minds" of the civilian population?[15] It is a debate that predates not only the conflicts in Iraq and Afghanistan, but also the US's earlier extended adventure in counterinsurgency in Vietnam. There is evidence, for example, of the British thinking very much along population–centric lines about the insurgency now known as the American revolution (Ricks, 2008). The debate is also anticipated by the debates surrounding "strategic bombing" in the first half of the 20th century (van Creveld, 2011).[16] That this debate has persisted for so long attests first and foremost to the complexity of warfare where civilians are key participants, and how this complexity contributes to the difficulty involved in determining practical courses of action where they are.

Some of the difficulty in making sense of insurgency and counterinsurgency can be attributed to the meaninglessness in this context of many of simplifying assumptions that allow those of us who study war to make some

[14]To the best of my knowledge, the terms are due to David Kilcullen, (Kilcullen, 2009)

[15]See chapter 5 for a more elaborate discussion of this distinction and its implications.

[16]See chapter 5 for more on the American Revolutionary War and strategic bombing, and their relevance to the current project.

sense of it. Most obviously, the "black boxing" of the internal politics of states advocated by Realists and their fellow travelers, on the grounds that there is always an objective and easily identifiable set of "national interests" that all states ultimately pursue, cannot be brought to bear on the study of insurgency and counterinsurgency:[17] As already suggested above, insurgents and counterinsurgents cannot necessarily claim or rely on the support of even a plurality of the population they purport to represent. The attendant multiplication of relevant actors, goals, and actions makes even the identification of a discrete list of potential explananda and explicantia a challenging task, and the identification of clear causal relationships among them even more difficult. Even the seemingly straightforward categories "insurgents" and "counterinsurgents" are complicated immensely once we acknowledge the extent to which both sides are reliant upon the often mercurial loyalty and support of people from the therefore equally fuzzy category "civilian populations."

Given these complications, one way of understanding the disagreement between advocates of population– and enemy–centric counterinsurgency is as a disagreement over the extent to which insurgents and counterinsurgents can be likened to states in a conventional war setting. That is, to what extent do insurgents and counterinsurgents face problems of collective action and coordination, and when they do, then what are the implications of these problems? The population–centric approach begins with the assumption that insurgent

[17]And there is increasingly good reason to believe that this approach can do little to illuminate even conventional wars, as I discuss in the next chapter.

organizations face significant problems of both coordination and collective action, and that a successful approach to counterinsurgency should seek to exploit those problems. The enemy-centric approach, on the other hand, takes the insurgent organization more or less as given, and seeks primarily to find and kill its members. Put another way, the debate is one over what costs are more important — the direct, material costs of war, or the political costs that flow or potentially, and perhaps more importantly, *blowback* from them — and over *how* and *when* they are important.

In addition to confusion and disagreement over the question of what are its possible causes and effects, there is the possibility in studying any war of *confusing* cause and effect, or of failing to properly account for the potential interdependence of phenomena, and again this problem is even more acute when it comes to the study of insurgency and counterinsurgency. The multiplication of categories of relevant actors in non-conventional warfare makes the already seemingly Sisyphean task of finding some "unmoved mover" behind an event even more difficult. It is not merely that any potential cause may in fact be merely an epiphenomenon — i.e. a phenomenon related to the outcome of interest not because it is a cause, but because it has a cause in common with it. It is also that at any point in time a phenomenon could be either a cause or an effect. Returning to the three fuzzy categories of relevant actors in an insurgency — insurgents, counterinsurgents, and civilians — it is not only that the insurgents' and counterinsurgents' behavior is influenced by the support of civilian populations, but that civilian populations will offer or deny

that support to insurgents and counterinsurgents based on the insurgents' and counterinsurgents' past behavior.

Thus, another way to think about the disagreement between advocates of enemy– and population–centric counterinsurgency is that it hinges on the question of how to calibrate violence to ensure the greatest possible attenuation of either the insurgents' forces, the insurgents' base of support within the civilian population, or both. Advocates of population–centric counterinsurgency might worry that too great an emphasis on killing might paradoxically add to insurgent organizations' strength by increasing the likelihood of collateral casualties, and therefore turning civilians against the counterinsurgents and increasing their willingness to support and even join forces with insurgents. The recruitment opportunities for insurgents provided by an improperly calibrated battle might then exceed the casualties they suffered while fighting. Conversely, advocates of enemy–centric counterinsurgency might worry that too little an emphasis on killing potentially allows insurgent organizations to recruit new members faster than the counterinsurgents can kill them. Thus the insurgents grow in strength while the counterinsurgents are preoccupied with policing the civilian population.

While it is often useful to assume otherwise, similar problems also attend the analysis of the behavior of states. It may well be that states can take for granted the support of their citizens and/or subjects most of the time, but it is important that we think very carefully about why that is, under what circumstances states might lose that support, what happens when they do,

and what states might do to avoid that outcome for themselves or precipitate it for others. As I will suggest in the next section, answers to these questions can go a long way towards explaining states' behavior in conflict settings.

1.4 The first Gulf War

During the run-up to Iraq's invasion of Kuwait in 1990, in a discussion of his territorial ambitions and his beliefs concerning the United States' willingness to obstruct them, Saddam Hussein remarked to U.S. Ambassador April Glaspie: "Yours is a society that cannot accept ten thousand dead in one battle" (Stein, 1992). A number of scholars of international relations have cited this quote as evidence of a head of state perceiving some difference in military resolve between democracies and other kinds of states, and as the beginning of an explanation of some states' willingness to engage even militarily superior democracies in conflict (Mueller, 1994; Reiter and Stam, 2002).

Hussein's statement appears to evince the belief that democracies (or, at the very least, America's particular form of democracy) are more sensitive to non-civilian war casualties than other kinds of states, and furthermore that this sensitivity puts them at a disadvantage in many kinds of militarized disputes. If Hussein and other world leaders hold such beliefs — regardless of whether these beliefs are true — then these beliefs should have some effect upon such leaders' behavior towards the United States and other democracies.

It may be difficult to imagine, given current technologies and norms of tactical planning, that the United States would ever suffer ten thousand deaths

in a single battle, especially in the aftermath of Hussein's armies' spectacular failure on two occasions to inflict upon American forces anything close to that number.[18] It is, however, possible to imagine the United States, through a process of attrition, suffering ten thousand or a great deal more deaths in a protracted war. If Hussein's statement was to be taken seriously as anything more than mere bluff and bluster, it was as an acknowledgment of the effects that the American voting public's response to such a process of attrition might have had on the United States' ability to persevere and prevail in a war against Iraq. If we wanted to imagine Hussein a particularly apt strategic thinker, we might furthermore expect him to have also considered that the United States' possible sensitivity to attrition could bring about a change in its tactics or strategy, and that this change might in turn affect the rate of attrition, and so on, and that anticipation of such behavior (as well as his potential anticipation of the United States anticipating such anticipation, and so on) would influence his attitudes and actions towards the United States.

Putting aside for the moment questions about U.S.–Iraq relations, the strengths and failings of Saddam Hussein's strategic intellect, or the particular phenomenon of casualty sensitivity, we might reasonably extrapolate from this brief development of the implications of Hussein's remark some expectations about the intersection of domestic and international politics in general. First

[18]That said, the development of these technologies and norms is almost certainly in large part a reaction to, and therefore evidence of, the very casualty sensitivity to which Hussein was purportedly alluding, as Goff and Tollison (1987) and Schornig and Lembcke (2006), among others, suggest.

of all, domestic politics should have some effect on international outcomes: Regardless of whether, as some Realists continue to argue, even democratic states pursue their own clearly and objectively definable security interests independent of internal politics, there is at least anecdotal evidence that other states might occasionally look to these internal politics for clues about their adversary's future actions, and choose their own actions accordingly. Thus, even if US domestic politics do not directly affect US foreign policy, they may very well have an indirect effect via their influence on other foreign policy leaders. Second, many of the domestic political phenomena that might influence the course of a given war will be dynamic in nature. That is, even when political institutions remain static, the domestic political phenomena that occur within the context of these institutions will exhibit some forms of change over time, and these changes will affect international outcomes. In the example above, it is not merely the fact that Hussein could observe the internal politics of the United States that might potentially drive some of his decision-making, but rather specifically what he might have observed — fluctuations in public opinion regarding the war, caused by increasing war casualties. Third, these phenomena will exhibit what econometricians call simultaneity or multiple endogeneity: That is, they will in turn be affected by the phenomena they have affected (and will in turn affect those phenomena again, and so on). We expect Hussein's decisions to change according to changes in public opinion, but also that public opinion will change according to outcomes reflecting changes in Hussein's decision-making (in fact, that Hussein's actions might affect Amer-

ican public opinion would appear to be the whole point of his statement to Glaspie).

Since about the mid–1990s, the first of these propositions has become less controversial, as the Realist "black-boxing" of domestic politics has ceased to dominate the international relations literature. Propositions two and three, however, while never exactly controversial, have played a strikingly small role in the theoretical and empirical literature that has taken up the task of opening the black box of domestic politics. In fact, failure to account for dynamics and simultaneity among political processes is endemic in most all of the empirical and theoretical literature on international relations, even where domestic politics are not considered.[19] The approach I propose explicitly addresses these shortcomings, by employing models that allow for measures of change over time and simultaneity of causality among domestic and international political phenomena.

One important addition to the above discussion is worth making at this point: There is no reason to imagine that Hussein was immune to the domestic political pressures he attributed to the United States. Thus the three implications of his statements about the United States should apply equally to him and to any other political leader, including the leaders of insurgencies. Whether and to what extent political leaders can count on the support of the population they claim to represent, and what are the greater implications of

[19] Although the recent literature on war termination and "bargaining while fighting" has taken great strides in addressing some aspects of this failure.

that support (or lack thereof), are often open questions — they are, in fact, often one of the reason belligerents fight in the first place.

1.5 Structural outline

Broadly speaking, what follows is an exploration of the role of domestic politics in conflict bargaining if disagreement leading to war is due to differences in *beliefs* rather than *information*. Smith and Stam have already noted some implications of this change in perspective for understanding conventional, bilateral conflict settings. But what of irregular warfare and/or conflict settings with more than two relevant actors? As it turns out, characterizing disagreement in such settings as a matter of differences in beliefs implies quite different patterns of behavior and outcomes than does characterizing disagreement as a matter of differences in available information, and the former characterization fits much better with empirical reality than does the latter.

I explain my research methods in more detail in chapter 3, but I offer here a brief discussion of the approach I have chosen and my reasons for doing so, as a means of explaining the structure of the following study.

My goal has been to think very carefully about how to measure the *interdependence* of conflict related phenomena *over time*. This has led me to explore and utilize empirical models that allow for what econometricians call "simultaneity," in particular vector autoregression models. These models allow for somewhat less specificity than usual in connecting empirical and theoretical

models, such that they are often accused of being essentially "atheoretical." While there is some truth to such charges, insofar as these models do not allow for the full identification of all of their structural parameters, for the purposes of my project, this lack of specificity is actually ideal. As I explain in chapter 3, my use of empirical analysis in this project is essentially *exploratory*, aimed at the goal of building, rather than confirming, a theory.

Thus, while I have made a serious attempt to link my empirical and theoretical models, my approach to doing so has been somewhat different from what is becoming the common practice in political science. The EITM — empirical implications of theoretical models — movement has challenged political science researchers to be more careful in linking their theories to the methods they use to test them, focusing on the usefulness of mathematical models in ensuring that the structural parameters of their empirical models are properly identified measures of the causal mechanism suggested by their theory (Granato and Scioli, 2004). While this is indeed a noble cause, my own project is more accurately understood as an exercise not in theory testing, but in theory building, and thus I take an approach almost exactly opposite that suggested by EITM. My focus has been on discovering not the empirical implications of a theoretical model, but rather the theoretical implications of an empirical model.

The theoretical framework developed herein is the end result, rather than the beginning, of this research project. My theory has its origins, in part, in two of the *negative* results of analyzing data I have collected: While many —

as the discussion above suggests, not only academics such as Kenneth Schultz, but also prominent political actors, such as Saddam Hussein and, as I discuss in chapter 4, George W. Bush — believe that public opinion in a democratic state can contain valuable (or, from that state's perspective, potentially harmful) information about that state's resolve in a time of war, as it turns out there is no evidence that insurgents in Iraq pay any attention whatsoever to American public opinion. While this is not definitive evidence that insurgents *ignore* American public opinion, it strongly suggests that they might. Furthermore, they may be right to do so, as the American public appears to have little awareness of the realities of the war — short-term fluctuations in American opinion are unrelated to any actual changes in the progression of the war itself — and are therefore a poor mechanism for the discovery of any useful information about America's capabilities and/or resolve.[20] These negative results suggested the need for an alternative model for analyzing the role of domestic politics and mutual optimism in the war in Iraq, and are one of the fundamental assumptions underlying the theoretical model I have developed to explain that role. While this approach to theory building is certainly not unique, transparent accounts of the process — which I have endeavored to provide in what follows — are quite uncommon. Chapter 2 provides more details of this process of inquiry.

In chapter 4, I begin by applying a variant of the "audience costs"

[20]Long run trends, both in the conflict itself and in American public opinion, are another story. See chapter 4 for more on this topic.

models of the relationship between domestic and international politics to the war in Iraq. Here, I follow the extant literature in assuming that an *informed* public can hold its foreign policy leaders *accountable* for their decisions, and thus making their decisions more *credible* to foreign adversaries (Schultz, 1998, 1999, 2001a; Brandt and Colaresi, 2008). In the Iraq war, then, I expect to see evidence of the American public paying attention and responding to the war, of American foreign policy leaders paying attention and responding to the American public and Iraqi insurgents, and of Iraqi insurgents paying attention and responding to the American public and American foreign policy leaders. But what I find are the results already mentioned above: The general American public appears to be completely uninformed about the realities of the war. Insurgents pay little to no attention to this uninformed public opinion. American counterinsurgents, however, are surprisingly responsive to public opinion.

Chapter 5 describes the end result of this project, an attempt to both explain and build upon the findings in chapter 4: a bargaining model of conflict that allows for belligerents to learn, in the short–run, the wrong lessons from fighting. Where dynamics such as those described in chapter 5 obtain, belligerents initially become more rather than less mutually optimistic as they fight, potentially both increasing levels of hostility among them and protracting the conflict. In the long–run, expectations and beliefs converge, just as they do in other models of bargaining and fighting to overcome mutual optimism.

In the conclusion, I summarize my results and my argument. I end

with a discussion of the limitations of the current project and directions for future research.

Chapter 2

Literature review

The present project is an attempt to address a number of related theoretical and empirical puzzles that the rationalist literature on war has not sufficiently addressed: What factors might cause/allow belligerents to change (or, conversely, to *not* change) their tactics or strategies during a war? What factors might cause/allow fighting to end? What factors cause/allow some wars to last so much longer than others? In attempting an answer these questions, I am contributing to and synthesizing two distinct rationalist literatures. The first of these focuses on audience costs, and in particular *domestic* audience costs — i.e. the capacity of political masses to punish political elites that make decisions counter to their interests[1] — and how these costs and their manipulation by belligerents affect conflict outcomes. The second is the literature on bargaining while fighting, especially the subset of that literature focusing on the phenomenon of "mutual optimism" — i.e. when the sum of belligerents' expected probabilities of winning is greater than one[2] — and how war might be a mechanism for overcoming it. I am also proposing an alternative to the

[1] One might also consider analogous "audience rewards" — i.e. the capacity of political masses to recompense political elites that make decisions consistent with their interests.

[2] See below for a more elaborate explanation of this phenomenon.

subset of the war termination literature that suggests that protracted wars are due to the difficulty states face in making credible commitments not to violate post–war peace settlements. A brief survey of each of these literatures will thus serve as a necessary preface to an elaboration of the novel features introduced by this project.

2.1 Moving beyond the Realist view of domestic politics

It is first of all important to acknowledge the historical context out of which these literatures evolved. While they all mostly reflect the developing consensus among international relations theorists that the security interests of states are not always obvious, and that determining what those interests are should be one of the primary goals of any study of the behavior of states, until very recently the conventional wisdom in IR was quite different. In fact, throughout the Cold War, the dominant paradigm in American academia for understanding international politics held that the opposite was true. Claiming the mantle of Machiavelli's famous exhortation to describe the world as it is rather than as we wish it were (Machiavelli, 2005), Realists maintained that explanations and even predictions of international outcomes could be derived from narrowly understanding all states as unitary actors, pursuing their own selfish, security-driven interests in an anarchical, "self-help" system. Realists further assumed that states' security interests were objective, easily identifiable, and universal, and would necessarily trump any of the potentially competing interests of specific individuals or groups living within them (Waltz,

1979). During the latter years of the Cold War, Realists enjoyed enormous influence not only within academia, but within America's foreign policymaking community, with some self-described Realists occupying Presidential Cabinet positions. It is not an exaggeration to say that theirs was the dominant paradigm for thinking about international politics by both academics and policymakers in the US for many decades.

In recent years, Realism has hit upon hard times. Scholars are increasingly skeptical about the ability of Realism to describe or explain the reality of international politics. On the empirical side of things, a number of articles published in the last two decades have called into question whether the "structural" or "systemic" variables proposed by Realists are in fact the best predictors of the occurrence of alliances, trade, war, or any significant international political phenomenon (Gartner, 1998, contains a good review of the empirical literature questioning Realist arguments.). By itself, this empirical failure does not necessarily mean the doom of Realism as an approach to the study of international politics. Realism, after all, is a self-consciously and transparently reductive and stylized approach to thinking about interactions between states, and even if it does not accurately describe or predict how states do behave, it might have a great deal of utility in explaining how states would behave if its heroic simplifying assumptions were met, which could in turn be useful as a baseline against which real world phenomena could be compared. But authors such as Powell, Schroeder, Wagner, Lake, and Frieden have called into question even Realism's claims to be able to explain the abstract (and perhaps

non-existent) world of unitary actors interacting in a setting of anarchy, pointing out that Realism's conclusions do not follow from its premises (Wagner, 1993, 2000; Schroeder, 1992, 1994; Powell, 1991, 1994). As these authors and others have suggested, at a minimum there is missing from Realist accounts of states' behavior some mechanism which might cause states to *want* to pursue wealth and power beyond their already established borders — Realists' claims to the contrary, the structure of international anarchy does not itself contain or imply such a mechanism. As Wagner puts it, "if it were common knowledge that all states wanted only to maintain their independence, it is hard to see why interstate wars would occur. To explain the occurrence of war, therefore, the leaders of at least some states must believe that there is at least some probability that other states might in addition want something that could be acquired by the use of force, and therefore it is important to know what that might be and how information about such interests could be acquired" (Wagner, 2010). In the absence of such interests, or at the very least *beliefs* about such interests, there is no reason to expect the international system to be characterized by mutual distrust, and the perpetual state of insecurity that supposedly logically follows from international anarchy alone clearly does not. Explaining international conflict thus requires explaining why states have such interests in the first place, or why other states might believe that they do.

Despite these seemingly fatal criticisms, Realism is far from dead, but rather is in the midst of a process of self-reinvention. As evinced by the now infamous article (and subsequent book) by Realists John Mearsheimer and

Stephen Walt on the supposedly pernicious influence on American foreign policy of the "Israel lobby" (Mearsheimer and Walt, 2006), as well as the growing (related) literature from those scholars Gideon Rose has dubbed "neo-classical Realists" (Rose, 1998), Realism today, in an inversion of its Machiavellian roots, appears to be transforming itself from a set of descriptions and explanations into a set of normative prescriptions and warnings: Realists no longer necessarily take it as given that states *do* behave as unitary and purely self-interested actors, but only that they *should*, and they are increasingly willing to use domestic politics as an explanation for those instances when they *do not*. These authors have additionally put themselves in the rather strange position of arguing that domestic politics only explain the failure of states to act in their own best interest, while the usual Realist arsenal of explanatory variables — anarchy, the distribution of economic and military power, etc. — still, of course, account for the successes. But how can this be? If domestic politics can in some cases intervene to cause states to act against their own objective interests, then we must explain why it is that in other cases they fail to so intervene. If we use domestic politics to explain some international outcomes, then we must, if only through uncovering the reason for their absence, allow them to play some part in explaining all international outcomes. Or, put another way, does not the very fact that Mearsheimer and Walt disagree so strongly with those they identify as the "Israel lobby" about what America's foreign policy should be call into question the notion that America's "best interests" can be objectively and easily defined? As the quote from Wagner

above suggests, discovering what, in fact, determines states' interests and/or states' beliefs about other states' interests, and what those interests and beliefs imply, are among the primary tasks facing IR scholars. As it turns out, some of the better answers to these questions can only come from a close analysis of states' domestic politics.

The present project is in many ways profoundly indebted to Realism, in that it arose in part from a consideration of the implications of Realism's transformations in the post-Cold War era. Regardless of the validity of its arguments, Realism has had a significant influence on American foreign policy — Richard Nixon's and Gerald Ford's National Security Advisor and Secretary of State Henry Kissinger being perhaps the most famous of all Realists.[3] But so, for that matter, have other, non-Realist approaches to thinking about international politics, such as neo-conservatism under George W. Bush and liberal internationalism under Bill Clinton.[4] If American political elites' very ability to evaluate the success or failure of their actions hinges on the paradigmatic lens through which they view those actions — such that two well informed

[3]Though it is an open question whether and to what extent Kissinger would actually agree with the evaluations of Mearsheimer or even Waltz.

[4]The distinctions between these approaches are perhaps not as stark as the above suggests. This is due in large part to the difficulty inherent in mapping any scholarly, theoretical paradigm to a corresponding set of pragmatic policy implications. (Kupchan and Trubowitz, 2007), for example, classify Nixon's foreign policy as "soft-liberal interventionism" — and thus, perhaps, consistent with the foreign policy approaches of Clinton and others. Similarly, some have suggested the existence of commonalities between Bush's approach and the "liberal interventionism" of Woodrow Wilson (Smith, 2009). For my purposes, it is sufficient to note that each of these approaches to thinking about foreign policy implies at least a slightly different set of interpretations of international political phenomena.

groups of Americans can look at the same information about the same foreign policy and come to very different conclusions about its likelihood of success — shouldn't political elites from different states also be able to disagree in this manner? What would the implications of such a disagreement be? Answering this question requires taking seriously the role of internal politics in determining international political outcomes.

2.2 Domestic audience costs

The "other means" that Clauswitz famously suggested make war distinct from normal political intercourse for the most part take the form of the direct or indirect imposition of costs. Unlike other kinds of political and economic engagement (with the exception of those that involve the possibility of labor strikes), the failure to reach an agreement in an international crisis setting does not necessarily lead to an end to bargaining and the re-imposition or maintenance of the status quo and, at worst, a temporary delay or breakdown in the provision of some contract-dependent good. Rather, an international crisis setting allows for the possibility of the involved parties imposing strong penalties upon each other for their failure to reach an agreement, and then to force a further stage of bargaining. Under the "black box" models proposed by Realists, where the broader interests of the state are understood to be the same as the interests of the persons in charge of making foreign policy decisions, it seems necessary to take a rather narrow view of these costs. Losses of reputation, treasure, and lives should only matter insofar as they dimin-

ish a state's ability to fight and win future battles. But breaking open these black boxes to acknowledge the multiple actors whose interests potentially influence foreign policy decision making complicates how costs operate. At a minimum, we might expect some kinds of states — democratic, autocratic, mixed — to have lower thresholds of cost tolerance than others, as Saddam Hussein suggested was true of the United States and Iraq. *Ceteris paribus*, states with lower cost thresholds might have certain strategic advantages and disadvantages when compared to their more cost tolerant counterparts. But more interestingly, how different domestic audiences might react to these costs, and how these reactions might change or be changed over time, have serious implications for international conflict outcomes. As mentioned in the previous chapter, it might also be that material costs only matter insofar as they translate into political costs.

Loss of life and limb is the most obvious, and perhaps the most definitive, of the costs associated with war, and what follows will focus primarily on the direct and indirect consequences of different actors' sensitivity to war's casualties. The seminal study on the subject by Mueller (1973) suggests that we can expect American public support of a war to be a negative logarithmic function of American casualties: support declines throughout as casualties increase, but the decline is steeper at the beginning when there is a large number of weak supporters (i.e. "fence–sitters") who are easily discouraged by failure. Strikingly, he finds no relationship between the vocal-ness of opposition and the rate of declining support. That is, the rate of declining support due

to mounting casualties was the same for both wars, despite the much more publicly visible domestic opposition that the Vietnam war garnered.[5]

A number of effects at the international level have been proposed as following from this casualty sensitivity and other, similar attributes that are supposedly unique to democracies like the US. Democratic leaders, concerned with losing favor among their electorate due to the mounting costs in human lives related to war, will be more restrained in choosing their battles. They will more often choose to fight those wars that are most vital to national interests, and that they are most likely to win. Selection effects due to these choices will mean that democracies will, on the whole, win more wars, and do so more quickly and decisively, than other kinds of states. On the other hand, when democracies become involved in longer and more costly wars, then they are less likely to prevail than are other kinds of states (Reiter and Stam, 2002, contains a useful survey of this literature and its hypotheses).

The notion that democracies are directly hampered in attaining their strategic goals by domestic sensitivity to war casualties has been the subject of some dispute. Christopher Gelpi, Peter Feaver, and Jason Reifler (Gelpi et al., 2007; Feaver and Gelpi, 2005), echoing Eric Larson a decade before (Larson, 1996), have suggested that the United States suffers from casualty sensitivity only when casualties are accompanied by a waning of the voting public's belief that the war in question is necessary, just, and ultimately winnable. But even

[5]The significance of this last point will be elaborated in chapters 4 and 5.

Gelpi, Feaver, and Reifler's caveats take for granted that changes in a democracy's, or any state's, success on the battlefield can have some effect upon the willingness of the domestic public to support a war, and, more importantly, that the consequent fluctuations in support have serious implications for that state's ability to continue fighting.

Furthermore, as the above discussion of Hussein's remarks suggested, democracies are not the only kinds of political organization that can be affected by domestic audiences' sensitivity to the costs associated with war. This point is not merely hypothetical: As recent work by Hein Goemans (2000) has shown, the leaders of autocracies should also be concerned about war costs — perhaps more–so — because their internal enemies would be happy to use such costs as a pretense for their ouster, and the repercussions of being removed from power tend to be much more dire for autocrats than for democrats. Non–democracies should therefore also be sensitive to the domestic audience costs of their foreign policies.

But what are the mechanisms through which sensitivity to war costs translates into behavior? Building upon previous arguments by Schelling (1960), James Fearon has suggested a number of implications that follow from the public nature of international crisis bargaining, its observation by domestic audiences who are likely to directly experience the costs of a war resulting from a failure to reach a settlement, and the observation of those domestic audiences by belligerents (Fearon, 1994). He argues that as domestic audiences observe and assess the performance of leaders, it can become increasingly difficult for

the leaders to back down or make concessions, for fear of appearing weak. Thus more constrained a leader is from making frivolous threats by such audience costs, the more credible the threats she does make, thus allowing states that are so constrained to more easily resolve crises without resorting to war. Thus there is a clear connection between the costs of war and the behavior of belligerents.

Fearon's conception of audience costs does not specify a mechanism through which foreign adversaries become aware of whether and to what extent a leader is constrained in a given moment, such that the *ex post* costs domestic audiences can impose translate into *ex ante* strategic advantages or disadvantages. It therefore offers little in the way of practical explanation of any observed variations in states' behavior and in crisis bargaining outcomes. Kenneth Schultz's research addresses some of these shortcomings (Schultz, 1998, 1999, 2001a). Schultz argues that it is the existence of opposition parties in democracies that allows democracies to make more credible threats than non-democracies. Opposition parties are likely to support the party in power when it makes threats that it can actually carry out, so that they will not be on record as having opposed a successful and therefore popular war when the next election comes around. They will oppose the party in power when it makes threats that it cannot carry out, so that they will be on record as having opposed an unsuccessful and therefore unpopular war when the next election comes around. Thus threats made by democracies can be "confirmed" in a manner that threats made by states without such electoral apparatuses

cannot.

While Fearon and Schultz are primarily concerned with crisis bargaining, i.e., bargaining that takes place in the shadow of war as a prelude to (or means of avoiding) actual fighting,[6] their theories do contain some implications about the behavior of belligerents who are already at war. If political masses can and do punish their leaders for making decisions contrary to their interests in the lead–up to war, why wouldn't they similarly punish their leaders for making decisions contrary to their interests during the war itself? There are, however, some problems with the operationalization of the mechanisms proposed by Fearon and Schultz, especially when considering their implications for bargaining while fighting. First among these problems is the unjustified assumption that electorates would automatically punish their leaders for backing down after making a threat (Schultz, 2001b). Why would this be the case? There have been some heroic, but ultimately unsatisfying attempts to answer this question (Smith, 1998; Guisinger and Smith, 2002), the most convincing among them the suggestion that electorates hold the belief that their statesmen will do better in executing their foreign policies when they have a reputation for honesty, and when statesmen get caught in a bluff this reputation is compromised (Guisinger and Smith, 2002). But surely such a belief should not be universal and unequivocal. Might not there be situations where electorates would want their leaders to bluff or even to start wars without committing to

[6]This is true of much of the literature on "two–level games" that their work can be understood to be a subset of (Putnam et al., 1993).

fight them to the finish? And if there is some value to occasionally bluffing, then should not electorates expect or even want their leaders to take occasional risks with their reputations for honesty, and understand that these risks will not always pay off?

The developing literature on "bargaining while fighting" (see below) offers a similar challenge to the notion that domestic political audiences should automatically punish leaders for escalating and then backing down. In particular, Wagner's approach to thinking about war suggests that it can function as a means of information gathering (Wagner, 2000). Each escalation and each battle is essentially a means for states to acquire otherwise private and incommunicable information about the others' capabilities and willingness to continue fighting. In this context it makes little sense to expect domestic audiences to always punish leaders for essentially engaging in the gathering of vital intelligence.

Furthermore, assuming that states with reputations for honesty and steadfastness uniformly do better in the context of international conflict situations than do states with reputations for bluffing and backing down, it is unclear why domestic audiences should matter in the first place. If a reputation for honesty and trustworthiness is what ultimately determines international outcomes, and a desire to maintain such a reputation is what drives all states' behavior (as Anne Sartori (2005) and Andrew Kydd (2005) argue), then there seems little point in thinking about domestic politics — states should endeavor to avoid bluffing regardless of the feelings of their citizens or subjects — and

there would therefore appear to be very little to distinguish the behavior of democracies from that of non-democracies.[7]

For the purposes of the current study, another notable shortcoming of nearly all of the above approaches to thinking about audience costs is that they do little to explain the influence of domestic politics in *non*-democracies — except, perhaps, to (questionably) imply that they do not matter as much as they do in democracies. The "selectorate" theory of war proposed by Bruce Bueno de Mesquita and his various coauthors takes some steps towards addressing this shortcoming, generalizing the political competition with audience costs model to bring some shades of grey to the stark democracies–v.–non–democracies contrast common throughout the rest of the literature (Bueno de Mesquita and Siverson, 1995). They argue that the key to understanding the behavior of political leaders is the size of the winning coalition (relative to the number of enfranchised or otherwise politically empowered citizens) that brought them to office. Once they have gained office, these leaders will want to reward the people who helped them gain power, in order to ensure that they will also help them maintain it. When the winning coalition is small, the leader will pursue policies that are in the interests of this very narrow minority. As the winning coalition grows, their interests necessarily become broader and more varied, ultimately converging with the interests of the majority of the population at large, and the policies pursued by political leaders will like-

[7]Mearsheimer's recent work on the rarity of lying in international politics offers some confirmation of this intuition (Mearsheimer, 2011).

wise converge toward policies consistent with promoting some perceived public good. The authors argue that these calculations encourage democratic leaders (i.e. leaders with large winning coalitions) to pursue less risky and less costly foreign policy objectives, because decisions that turn out to be costly for the population as a whole could get them tossed out of office. Meanwhile, authoritarian leaders (i.e. leaders with small winning coalitions) will pursue more risky and costly alternatives: They can more easily offset any damage done to the public at large with selective rewards to the people that brought them (and whose continued support keeps them) in office. While this is a promising approach to thinking about the role of domestic politics, it is, like the other models thus far discussed, far too static, thus obscuring some of the more interesting implications of thinking of foreign policy in terms of gaining and maintaining winning coalitions. First of all, certainly the very size of the winning coalition is one of the things that a political leader's policies, or an opposition party's proposals, might be meant to influence.[8] Second, and perhaps more importantly for the present project, the *process* through which this winning coalition is formed and brings a given regime to power, and in fact *whether* the regime is ultimately able to take power at all, might be of some great importance, and are questions that Bueno de Mesquita's approach does

[8]I owe this last insight to discussions with Harrison Wagner. Note, however, that the desire to change the size of a winning coalition does not necessarily imply that political leaders' foreign policy decisions are directed towards *increasing* the size of said coalition. As E.E. Schattschneider's work suggests, a politician's success is a function of her ability to strike a balance between successfully *widening* her base of support and *limiting* the number of interest groups to whom she is beholden.

little to answer.

When explaining the behavior of belligerents in the context of an insurgency or civil war, these last points are especially important, as the answer to the question — "Whose will be the winning coalition?" — is precisely what they are fighting over.[9] Determining the implications of these points requires moving past two common features of "rationalist" theories of war — first, the "costly lottery" model of conflict, and second (and somewhat more controversially) the assumption of "common priors."

2.3 Bargaining while fighting: Mutual optimism and its remedy

Most approaches to thinking about war as either a form of bargaining, or the result of a failure in bargaining — including the models of Fearon, Schultz, and Bueno De Mesquita discussed above — typically treat war as a "costly lottery." In the typical costly lottery representation of conflict (Fearon, 1995), both parties have respective probabilities p and $1-p$ of winning the war, some value W they assign to winning, and they must pay some costs c_1 and c_2 to fight. Fighting such a war would thus have an expected value of $p \times W - c_1$ and $(1-p) \times W - c_2$.[10] Wagner remarks that such a "costly lottery"

[9] And as, for example, (Goemans, 2000) and (Trubowitz, 1998) work suggests, this may often be the case in interstate conflict as well.

[10] Note that I have simplified for the sake of space the usual version of a costly lottery somewhat, by conflating the share of the object of war each side receives with the subjective value that each side assigns to it. Under most circumstances, the logical ramifications of my simplified version will be the same as in the more complicated canonical model.

representation of war is equivalent to what Clauswitz calls "absolute war." That time is not a parameter means the cost of fighting and the probability of winning are constant, implying that they represent not a series of battles with variable outcomes, but rather a single fight fought to the finish. Thus waging war is essentially, in Clauswitz's words, "one decisive act," like the roll of a die (von Clausewitz, 1993, page 87).

In the absence of additional complications, states under these circumstances should always prefer a negotiated settlement to absolute war, because there are any number of mutually agreeable divisions of W that they would prefer to fighting (Fearon, 1995). State 1 should prefer any share larger than $p \times W - c_1$ to war, and state 2 should be willing to give state 1 any share less than $p \times W - c_2$ rather than go to war. Put another way, $c_1 + c_2$ represent the overall value lost by fighting — the inefficiencies of war — and both states should prefer finding some way of keeping and dividing up this value rather than fighting and losing it. The larger the values of c_1 and c_2 are, the larger the range of settlements that both states prefer to war, and as long as c_1 and c_2 are greater than zero — which we can assume given the loss of lives, territory, and treasure that war necessarily entails — some range of settlements will exist. Given this range of mutually agreeable divisions of W, why would an absolute war ever occur?

One answer to this question that has inspired much of the contemporary literature on conflict is mutual optimism (Fearon, 1995). It may be that one or both sides overestimate their probabilities of winning, assigning subjective

values to p — p_1 and $1 - p_2$ — such that $p_1 + 1 - p_2 > 1$ ($p_1 - p_2 > 0$). The larger are p_1 and $(1 - p_2)$, the smaller are the range of divisions of W that both would prefer to war. Once the difference between $p_1 \times W$ and $(1 - p_2) \times W$ is greater than $c_1 + c_2$, the range of possible settlements preferred to war disappears entirely, such that $p_1 \times W - c_1$and $(1 - p_2) \times W - c_2$ are greater than any division of W that the other side would agree to. Such mutual optimism on its own is not enough to lead to war. If mutual optimism were the only problem, states could avoid war by directly communicating to each other whatever factors they believe give them an advantage, thus disabusing each other of their respective incorrect assumptions about future success. In order for mutual optimism to lead to war, it must be the result of states having private information with an incentive to misrepresent, different beliefs about likely outcomes given the facts, or both.[11]

[11] As mentioned in the previous chapter, Mark Fey and Kristopher Ramsay claim to have shown that mutual optimism cannot lead to war (Fey and Ramsay, 2007). They argue that except when the negotiation of a possible settlement is characterized by "take it or leave it" offers (where "leaving it" automatically leads to war), or when the costs of negotiation approach the costs of war itself, the very process of negotiating and renegotiating the terms of possible settlements in the lead up to war allows for the unintentional communication of private information about relative strength. By making an offer, a state necessarily reveals to its adversary information about what it believes to be the probability of winning and the costs associated with fighting. If said adversary has an opportunity to make a counteroffer, then it will do so after revising its own beliefs in light of this signal, thus sending a signal of its own to the first state about these revised beliefs. The first state can then make a similar revision and make its own counteroffer, thus sending a new signal about its revised beliefs. And so on. As long as both states agree on the factors that contribute to the probability and cost of winning, then through the iterative process of making offers and counteroffers, and updating their beliefs in response to the other's most recent counteroffer, each state inexorably arrives at estimates of these probabilities and costs that are close enough to p for a range of mutually agreeable divisions of W to open up.

But in their attempt to "generalize" their model of mutual optimism, Fey and Ramsay

As Wagner suggests, there are two possible interpretations of "costly lottery" models. The first is that a choice between fighting and settling is "the result of a take-it-or-leave-it offer made by one state to another: if this offer is rejected then a fight to the finish will occur." But not all fights are to the finish — most, in fact, are not — because "a state might be willing to fight for a while in hopes of getting a better deal even though it would be willing to forego the possibility of agreement altogether," so a "costly lottery" model so understood can only explain a very small number of wars. The second is that such models "merely represent the prewar expectations of states concerning the terms of the final settlement and the costs that will be suffered before it is reached." But from whence these expectations come is among the most pressing questions we can ask about conflict, which "costly lottery" models can do nothing answer. Thus, Wagner suggests, despite their predominance, "costly lottery" models of conflict are of very limited usefulness.

Invoking Clausewitz's dictum that war is "the continuation of political intercourse, with the addition of other means" (von Clausewitz, 1993), Wagner

---

remove from their model all of the mechanisms that make mutual optimism possible in the first place. Thus it is no surprise that in their model mutual optimism cannot lead to war. As already mentioned above, the mutual optimism that can lead to war is either a result of private information with an incentive to misrepresent, different beliefs, or both. Fey and Ramsay have constructed a model where beliefs are assumed to be universally shared, and the communication of information is so frictionless that nothing can ever be private, regardless of the intentions of the actors. That such communication is rarely so frictionless is precisely what modeling conventions as "take it or leave it" offers is meant to represent — admittedly, rather clumsily. See (Slantchev, 2011) for more on these points. Given these objections, and that theirs is a costly lottery model of war that does not allow for non-common priors — modeling choices that largely drive their results — their critique has little bearing on the current project.

suggests that we should understand war — or, more specifically, what Clausewitz called "real war" — as a part of the bargaining process. Here both sides try to influence the terms of a negotiated settlement by fighting, with a more grandiose "total war" of disarmament as the usually never reached disagreement outcome. In other words, states are fighting not to win, but rather to determine who *would* win if they continued fighting.

This conception of mutual optimism suggests another a way it can lead to conflict that is not revealed when war is treated purely as a costly lottery. It need not be the case that both sides' overestimates of their likelihood of success be so great that there is no settlement they would both prefer to fighting (though this, of course, would also lead to fighting in Wagner's conceptualization). Rather, if one side believes it can by fighting reduce the other's estimate of the likelihood of winning, then it should expect to get greater concessions in the negotiated settlement that follows. Under these circumstances, it might choose to engage in a real war even if it expects it would lose in an absolute war. Thus conflict is possible even when one side has a clear material advantage, as is clearly the case in insurgencies such as the current wars in Iraq and Afghanistan.

In the context of Wagner's conceptualization, domestic politics should function in much the same way that they do in the "costly lottery" based "audience cost" models explored above. Here, the information revealed by domestic politics would presumably compliment the information revealed by fighting, thus facilitating the convergence of belligerents' beliefs and expecta-

tions about winning. In other words, when domestic politics come to bear, they can precipitate a war's termination. If this is the case, then domestic politics might tell us very little about protracted conflicts like the current war in Iraq — unless, that is, we are able to somehow measure the lower *salience* of domestic politics in such a conflict, and in so doing predict or explain its duration. As I will show in chapter 2, in the current war in Iraq, soon to be America's second longest military engagement (the war in Afghanistan is already in first place), it turns out that the domestic politics of the US have played only a very limited role, which may very well have had some relationship to its length.

Recent work by Smith and Stam offers a new approach to thinking about bargaining while fighting in general, and some interesting implications regarding the application of this concept to the context of insurgency and counterinsurgency specifically. As discussed briefly above (and in greater detail in chapter 5), Smith and Stam suggest that rather than following the convention in the rationalist literature of treating different beliefs as essentially *informational*, we can allow them to differ even when actors are given access to the same information. This requires eschewing the "common priors" assumption common to nearly all rational choice treatments of war (Harsanyi, 1967).

That belligerents may, as Smith and Stam suggest, "agree to disagree" raises the question of what they can agree to disagree about. In Smith and Stam's model, they can initially agree to disagree about their respective expected likelihoods of winning, leading them to war, and fighting allows them

to bring these expectations into alignment, leading them to a negotiated settlement. In chapter 5, I explore an alternative to their model wherein belligerents can disagree not only about their likelihoods of winning *before* fighting, but about who is in fact winning *while* fighting. I show that when counterinsurgents have strong beliefs about their own relative capabilities, but weak and contradictory beliefs about what approach to fighting — enemy– or population–centric — is most effective, fighting can cause their expectations about winning to *diverge* from rather than converge towards those of the insurgent. Furthermore, when these same conditions obtain, the counterinsurgent has the opportunity to learn which approach to fighting is, in fact, more effective, and change her approach if necessary.

2.4 Explaining protracted wars: Domestic audiences, mutual optimism, and credible commitments

In chapter 5 I show how a model of mutual optimism and learning while fighting can be adapted to account for, among other things, the protracted nature of some wars. Recent work by Dan Reiter suggests an alternative way of thinking about protracted conflicts, based on the "realist insight" that in the absence of a global government, there is no–one to enforce post–war settlements (Reiter, 2011). As a result, belligerents face credibility problems not only when it comes to their ability and willingness to start fighting, but also about their ability and willingness to stop fighting and keep to the terms of a treaty. If a belligerent's promise to honor a post–war settlement is not

credible, then fighting continues until it is.

Invoking Clausewitz, Reiter argues that there are therefore two kinds of war. The first, equivalent to Clausewitz's "limited wars," are due to what Reiter calls "informational" problems. These are wars fought due to mutual optimism, as described above and in chapter 5, and that end in a negotiated settlement once belligerents have overcome mutual optimism through fighting.[12] The second, equivalent to Clausewitz's "absolute" or "total wars," are due to belligerents' inability to credibly commit to a negotiated settlement, and typically end only when one side's capacity to continue fighting has been completely annihilated.

Reiter here sidesteps Clausewitz's assertion that limited wars are actually the only kind we ever observe, and that "absolute wars" are merely a "play of the imagination" — i.e. a thought experiment to help better clarify what real wars are (von Clausewitz, 1993). Instead, Reiter defines the concept somewhat more expansively than does Clausewitz, as a war that ends in one belligerent either annexing the other's territory, dismantling its military, or imposing on it a change of regime. With this definition in place, Reiter can say that there have been around 60 absolute wars since the early 19th century. Never mind that in many of these wars, fighting nonetheless continued long after the "absolute" victory of one of the belligerents — as was clearly the case

[12]Reiter does not concern himself with the intricacies of the Harsanyi doctrine, and does not distinguish between disagreements due to asymmetric information and disagreements due to differences in beliefs. Thus both kinds of conflict are, according to his rubric, "informational."

in America's invasion of Iraq. The inability of belligerents to make credible commitments to stop fighting might explain protracted conflict, but it isn't clear how what Reiter calls "absolute war" solves this problem and thus helps them to end fighting.

But as it turns out, the "commitment problem" identified by Reiter can explain very little, as it is based on an essentially false dichotomy between domestic and international politics — i.e., that in the former contracts are possible, due to the enforcement power of the state, but in the latter they are not possible, due to the absence of such a power. As Wagner suggests,

> such reasoning rests on a confusion between the government's role in enforcing agreements among individuals and the enforcement of the agreements that define the state itself. One such set of agreements defines the organization of the government, another its relation to its subjects, and a third the boundary between its territory and the territory of other states, and they are all subject to renegotiation by the use of force. There is no external enforcer of any of them, and therefore what enforces them all is a comparison of the benefits they provide with expectations about the consequences of trying to renegotiate them. A more economical way of saying the same thing is that all these agreements must be self–enforcing (Wagner, 2007).

While it is not clear how what Reiter calls "absolute wars" would help states reach self–enforcing agreements, what *Clausewitz* calls "absolute wars" might

serve quite well. It is probably the case that an actual absolute victory is not only impossible but, for victor and loser alike, actually *undesirable* given the effort that would be involved in realizing it. But coming to an agreement on who they *imagine* the victor of an absolute war might be is a necessary precondition of a self–enforcing agreement to stop fighting. Why, after all, would a belligerent keep fighting if she knew that another battle would likely only serve to bring her another step closer to disarmament?

But this brings us back to one of the central problems of the current project: Under what circumstances is it possible for belligerents to come to an agreement on who the likely victor of an absolute war would be? I take up this problem in chapter 5.

Chapter 3

Research design

3.1 The empirical implications of theoretical models

As mentioned in the previous chapter, the current project is organized in a manner that departs somewhat from many contemporary standards of political science research. In particular, it is a kind of inversion of the EITM (Empirical Implications of Theoretical Models) approach (Granato and Scioli, 2004). To explain and justify this inversion, I will briefly sketch the motivations behind the EITM movement and why I view my own work as essentially consistent with those motivations.

The original NSF–sponsored EITM workshop was organized primarily in response to the widely observed problem of non–cumulation in political science research (Achen, 2002), i.e., that after decades of using increasingly sophisticated mathematical models to make sense of political phenomena, political scientists have remarkably little to show for their efforts. Not only have political scientists not identified any "law–like" properties attending political phenomena, they have established very little in the way of widely accepted theories explaining what they have observed.

The EITM workshop participants suggested that a primary reason for

non–cumulation in political science research is political scientists' failure to properly link empirical and theoretical models. This failure, they argued, is due in large part to the increasing compartmentalization of theoretical and empirical research, precluding necessary dialog between the two. Thus, for example, when theoretical researchers develop formal models, they often do so with little or no basis in or even acknowledgement of the extent to which their foundational premises may be unsupported by empirical observation. Similarly, empirical researchers often treat peculiarities in their data as mere sources of statistical error to be corrected with statistical tools, ignoring the extent to which such peculiarities may be theoretically meaningful. Thus, without proper guidance from the other, neither is able to make any real progress.

Constructing bridges between empirical and theoretical research, according to the EITM approach, entails the use of formal modeling techniques to identify the structural parameters to be estimated with an empirical model. But basing empirical models on analogous formal models requires that formal models with clearly specified, observable, and testable empirical implications exist. For projects involving the study of electoral politics and, perhaps, comparative and international political economies, such models may indeed be in plentiful supply. But unfortunately, as R. Harrison Wagner has suggested, in projects involving the study of international conflict such models are in limited supply. By Wagner's reckoning, the problem is not so much that empirical researchers are not paying sufficient attention to the insights coming from the

theoretical literature, but rather that the theoretical literature has surprisingly little guidance to give empirical research. The vast majority of theories in political science treatments of international relations, Wagner suggests, are essentially incomplete. Not necessarily wrong (or, as more scathingly put by physicist Wolfgang Pauli, "not even wrong"), these arguments are nonetheless logically invalid, their conclusions not of necessity following from their premises.

If theoretical approaches to understanding international politics are largely made up of logically invalid arguments, this poses a formidable challenge to any researcher who wants to employ the EITM approach: It is not possible to derive clear empirical implications from a logically invalid argument. Any test of such an argument will be at best merely a test of any and all arguments that have similar premises. This, Wagner suggests, is one reason why the "debate" between Realists and Neoliberal Institutionalists has persisted for so long. Because both approaches seem to base their contradictory conclusions on nearly identical premises, it is not clear how any empirical test might decide between them.

3.2 The theoretical implications of empirical models

In fact, the situation may be even worse than Wagner suggests. The dearth of logically valid arguments in the study of international politics may be due in part to the essentially *cognitive* difficulties attending their construction, as Wagner suggests. But it may be due as well to the lack of well established

premises with which to construct logically valid arguments. With the possible exception of the poorly explained and understood "democratic peace," there are few empirical generalities that IR scholars agree upon.

Due in large part to the extent to which the agenda of IR research has been historically set by Realists, what is missing from most explanations of the behavior of states and other political organizations that participate in international politics is some notion of what it is — besides some vaguely defined interest in "security maximization" — that such organizations (or, more accurately, their members) want, and why they want it. These basic, underlying motivations, far from given, are among the most poorly understood and therefore important objects of inquiry in the study of international politics.

The problem is that there are few good places to start in determining what these organizations' motivations are or where they come from. As discussed in the previous chapter, in the rational choice literature on conflict bargaining, the most common approach to thinking about these motivations is to assume that they are determined in large part by costs that the organizations' base of support — the electorate in democracies and the "selectorate" in non–democracies — can impose on them. Kenneth Schultz and other scholars argue that because these domestic audiences can impose such costs on political elites, their actions can be quite revealing, communicating to rival organizations otherwise private information about their intentions and resolve. But as I suggested in the previous chapter, there are a number of serious shortcomings to such a conceptualization of the role of domestic politics in a conflict

setting, and therefore good reasons to imagine that we will be often unlikely to observe the specific patterns of behavior Schultz suggests follow from it.

The purpose of the current project, then, is to find a better model of this relationship between international and domestic politics by observing the realities of a specific conflict setting. To make sense of the war in Iraq, I begin with the "audience costs" model, and then see to what extent empirical realities conform to what is predicted by this model. The role of empirical analysis in this project is not to confirm or disconfirm the "audience costs" model. On the contrary, given the concerns expressed in the previous chapter, I begin with the assumption that there will be a mismatch between what the model suggests and what I actually observe. The point of empirical analysis, then, is to determine what the nature of that mismatch is, so as to construct an alternative model (see below).

3.3 Data

In order to more carefully evaluate the claims of the "audience costs" literature [1], I use vector autoregression (VAR) models to examine the dynamics of the insurgency and counterinsurgency in Iraq. I will explain VAR models in greater detail in the next section. But in this section I will discuss briefly the kinds of data required by such models, then describe and explain the data I collected for analysis and some of their peculiarities.

[1] As well as the rhetoric of US foreign policy elites; see chapter 3.

3.3.1 General requirements of the data

The analysis of time series data in general, and vector autoregression in particular, requires that data be *stationary*. A stationary time series is a stochastic process whose "mean and variance are constant over time" and for whom "the value of covariance between two time periods depends only on the distance or lag between the two time periods and not the actual time at which the covariance is computed" (Gujarati, 2002).

Most time series data are non–stationary (Gujarati, 2002). In practice, therefore, the analysis of time series data requires that these non–stationary data be transformed to stationary analogs before they can be fit to a statistical model. There are two approaches I take to effecting such transformations, depending on the source of non–stationarity.

The first potential cause of non–stationarity is that the data actually represent a number of distinct phases, each with its own mean, variance, and covariance between different time periods within the phase. The war in Iraq can be divided into such phases, i.e. the invasion, the initial occupation and insurgency, "the awakening," "the surge," etc. Combining data for each phase into a single, long time series would thus confound inference, as the resulting time series would be non–stationary. In analyzing the conflict, therefore, it is necessary to either employ empirical models that account for changes in "regimes," or to analyze each phase separately. I have chosen to take the latter approach, and limit the scope of my analysis to the longest such phase — beginning with the end of the initial invasion (marked by Bush's "Mission

Accomplished" speech) and ending before the commencement of "the surge."

The second potential cause of non–stationarity is the presence in the data of some kind of trend. A variable that more or less inexorably increases or decreases over time will not, needless to say, have a constant mean over time. Where more than one time series has such a trend, it is sometimes possible that a linear combination of the trending series will be stationary (these are called "cointegrated" time series). This is both convenient, allowing for time series analysis of the data without radically transforming them, and substantively meaningful, suggesting a long–run relationship between the variables. Otherwise, it is necessary to transform the data, usually by differencing, until the trend goes away.

3.3.2 Public opinion

To determine the effect of changes in foreign policy opinion on the behavior of US and Iraqi forces, I have collected into a time series polling data on public approval of Bush's handling of foreign affairs. Using the iPoll database at the University of Connecticut's Roper Center, I pooled data from polls on American public opinion regarding Bush's handling of foreign policy conducted by NBC News/Wall Street Journal, CBS News/New York Times, Time/CNN/Harris Interactive, Democracy Corps, Gallup/USA Today, Princeton Survey Research Associates/Newsweek, Pew News Interest Index, IPSOS-Reid/Cook Political Report, and Quinnipiac University. I used only polls with comparable questions and samples (Roper Center, 2007).

These polls were conducted at irregularly spaced intervals. For the purpose of time series analysis of these data, it was necessary to create a temporally consistent analog. In order to be consistent with the conventions of the empirical literature on reciprocity (Goldstein and Freeman, 1990), I need a weekly analog of the data. Rather than a linear interpolation, where weekly values are generated by taking weighted averages of values at the nearest preceding and succeeding dates, I use a constant interpolation, where each weekly value is set as equal to the value at the nearest preceding date for which polling data is available. I believe this constant interpolation makes the most substantive sense: The measures of public opinion published by these polling agencies are the only measures of public opinion potentially available to all of the relevant actors, and while there might be some gradual, linear change in public opinion occurring between polls, most of the relevant actors will not know about the direction or magnitude of this change until after the next poll has been published. Figure 3.1 presents the time series created using the constant interpolation.

There is an obvious downward trend in these data. This is easily apparent from merely eyeballing the time series plot of the data, but an augmented Dickey Fuller test of the time series data confirms its non–stationarity. As explained above, such trends interfere with accurate time series analysis. Thus, I have created a second time series of de–trended (first differenced) data, which can be substantively interpreted as the weekly change in public opinion regarding Bush's handling of foreign policy. Repeating the augmented Dickey

Fuller test on the de–trended data indicates that doing so has eliminated its non-stationarity. These de–trended data are in figure 3.2.

As suggested by the discussion above, using de–trended data is somewhat problematic: The elimination of the most obvious characteristic of a given variable (its downward slope over time) could also potentially eliminate whatever effect might have on the other variables. This would certainly be the case if the other variables of interest exhibited their own long–run dynamics. But as I show below, neither shows any indication of non-stationarity; all of their dynamics are short–run, and it makes little substantive sense that a long–run trend in one variable could explain short-run variations in another. De–trending the public opinion time series ensures that the focus of the statistical analysis is on these short–run dynamics. That said, the long–run, downward trend in American public opinion is far from unimportant. I will discuss its implications in the conclusion of chapter 3, and in greater depth in chapter 4.

3.3.3 Hostility

To see how belligerents react to changes in public opinion, as well as to each other, I need a measure of each sides' level of hostility toward the other over time. To create such a measure, I have used machine coded event data, relying upon Philip Schrodt's Tabari engine to code newswire feeds from the Agence France Presse wire service (Schrodt, 2006).

Tabari looks at wire service articles' lead sentences (which, by conven-

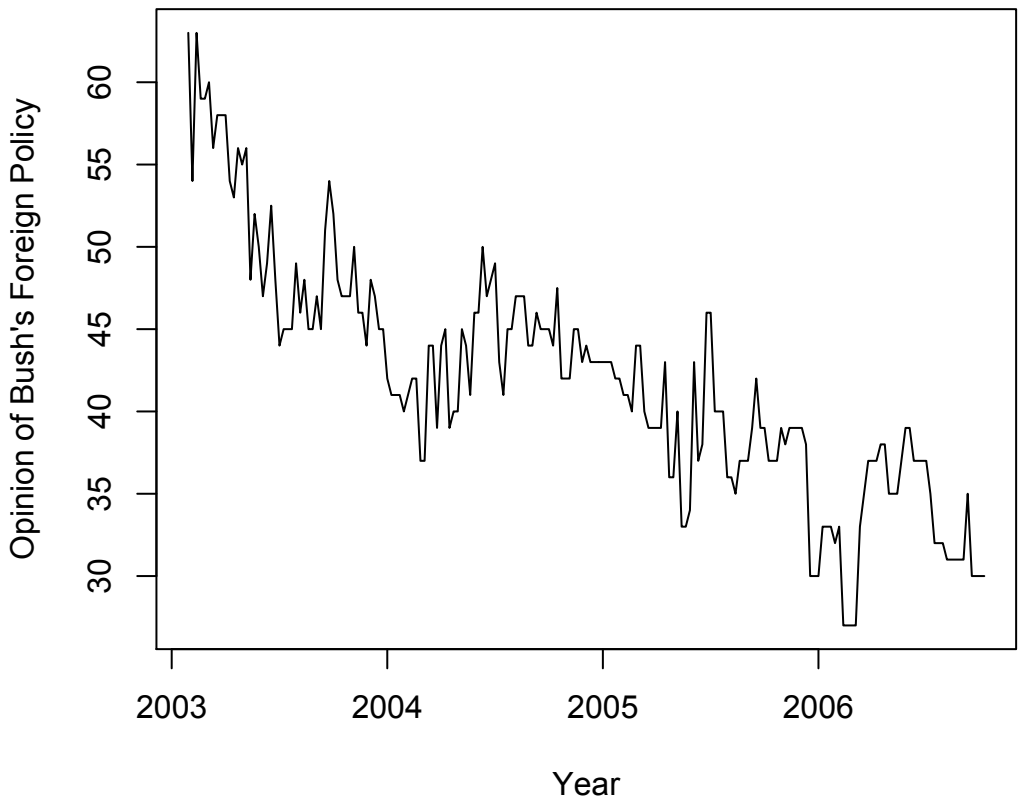


Figure 3.1: Regular intervals, constant approximation

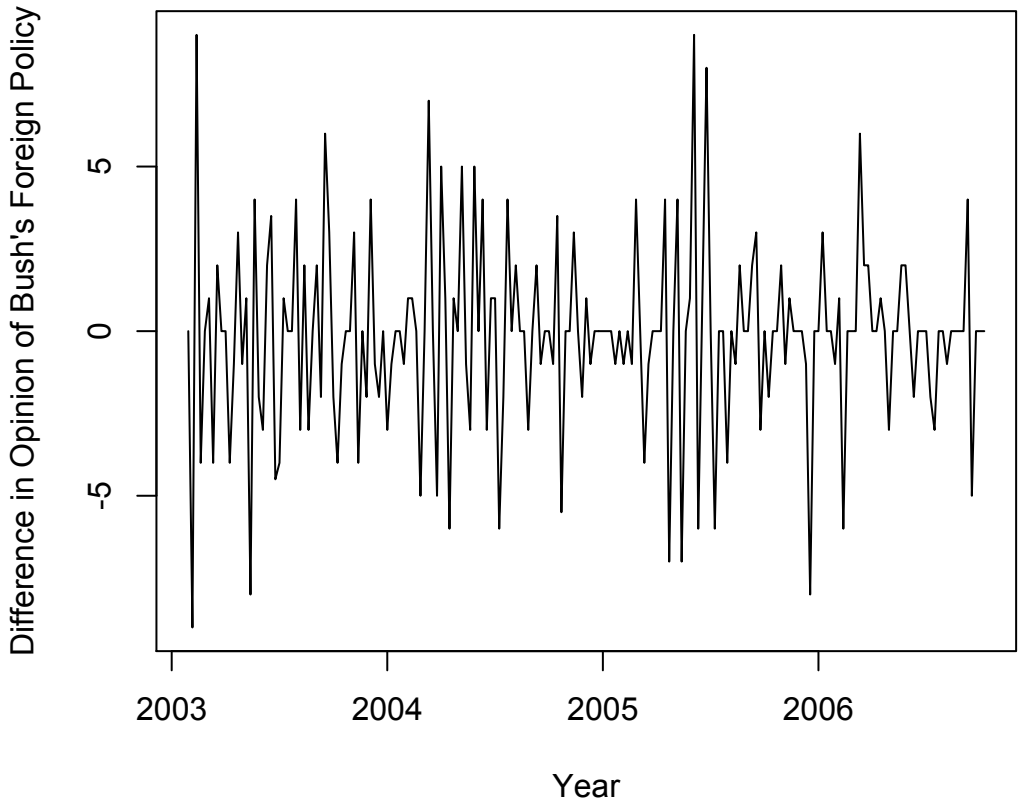


Figure 3.2: Public Opinion

tion, contain a declarative summary of the entire article), and tries to determine who the main actors are and what the lead suggests they've done to each other, based on various contextual cues and a set of dictionaries developed by the user. The main difficulty in using Tabari, besides the time-consuming process of downloading the enormous number of articles for it to read, is the development of these dictionaries.

The two primary dictionaries that Tabari uses are an actors dictionary and a verbs dictionary. The first of these identifies all of the actors of interest and associates each of them with a code representing the primary group to which they belong. In my project, I associated the names (and common variants thereof) of people who have fought on behalf of the United States (from actors as specific as "Lynndie England" and "Patrick Tillman" to the more generic "US soldiers") and those in the military or civilian government who have authority over them (again, as specific as "George Bush" or "Donald Rumsfeld" and as generic as "The White house," "The Pentagon," or even "The US") under the group USA. Then, I put the names (and variants thereof) of members of Al Qaeda in Iraq, the Ba'ath Party, and anyone identified as a Shiite radical, a Sunni radical, or an insurgent into the group IRQ (again, using both very specific and very generic terms to identify such actors). There are, of course, problems with lumping all of these groups together. Primarily among them, in the case of the group IRQ, is that many of the subgroups are fighting each other. However, the interdependence of public opinion and hostilities between the US counterinsurgents and Iraqi insurgents requires only

that problems of collective action and/or coordination be resolved *within*, not *among*, Iraq's rival factions.[2] My primary goal in creating my actors dictionary was to be as expansive as possible, but without including co–nationals who are not actually directly associated with or responsible for either set of belligerents. For example, I would not want Nancy Pelosi or John Kerry to be included under USA, despite the fact that they are, obviously, members of the US government, as Tabari would be unable to handle appropriately their vocal criticisms of the Bush administration. Nor would I want ostensible US allies such as Ayad Allawi or Nuri al-Maliki included under IRQ, for similar reasons. This turned out to be the most difficult part of using Tabari for the purpose of exploring the dynamics of insurgency and counterinsurgency: Even with a very sophisticated and thorough actors dictionary, it is too often the case that Tabari will incorrectly interpret, e.g., a conciliatory statement from the United States government to an already vaguely sympathetic post-invasion Iraqi government official, or vice versa, as a reduction in hostility between belligerents.

I created my initial actors dictionary by referencing existing dictionaries created by Philip Schrodt in preparation for his most recent research on general levels of conflict in the Middle East. I then expanded these dictionaries by including names of slightly more obscure actors in the US and Iraq, whose inclusion would have had only marginal effects (compared to the effort required

[2]Whether and to what extent these groups were capable of cooperating among themselves, and what the implications of such cooperation (or lack thereof) might be, are all certainly interesting questions, but are beyond the scope of this project.

to enumerate all of them) on Schrodt's analysis of the larger region, but which nonetheless would significantly improve the number and accuracy of identified events pertaining only to the conflict in Iraq. To give two examples, there was only one phrase in Schrodt's dictionary that would have identified Moqtada al-Sadr and the Mahdi army as key actors, despite the enormous number of permutations of phrases (e.g. 'Mehid', etc.) in AFP articles referring to these people, and neither Jay Garner nor Paul Bremer were included in Schrodt's dictionary at all. I further expanded the dictionaries by including actors who have gained prominence since Schrodt's study ended (such as Robert Gates). Finally, I expanded the actors dictionary by using a Perl text processing script to find and count all instances of phrases consisting of words with capital first letters and used the most frequently appearing words of the resulting lists to find phrases identifying relevant actors that neither I nor Schrodt had considered.

The verbs dictionary associates verb phrases with a numeric code indicating a slightly more generic action. Among studies of conflict employing event data, there are two now standard coding systems: WEIS, originally developed by Charles McClelland and published in 1984, and CAMEO, developed by Schrodt. The CAMEO system is more thorough, more coherent, and, according to Schrodt, more compatible with Tabari, so I elected to use this system. CAMEO consists of 20 categories, each assigned a two-digit number:

- 01: Make Public Statement

- 02: Appeal
- 03: Agree
- 04: Consult
- 05: Support Diplomatically
- 06: Cooperate
- 07: Provide
- 08: Yield
- 09: Investigate
- 10: Demand
- 11: Disapprove
- 12: Reject
- 13: Threaten
- 14: Protest
- 15: Exhibit Military Posture
- 16: Reduce Relations
- 17: Coerce
- 18: Assault

- 19: Fight
- 20: Attack with Weapons of Mass Destruction

Within each category, there are a number of sub-categories. For example, 19: Fight has the following sub-categories:

- 190: Use conventional military force, not specified below
- 191: Impose blockade, restrict movement
- 192: Occupy territory
- 193: Fight with small arms and light weapons
- 194: Fight with artillery and tanks
- 195: Employ aerial weapons

Tabari assigns to each lead it identifies as an event (i.e., a lead containing two actors and a verb phrase that it recognizes) a three or four digit number representing one of these sub-categories, based on associations established in the verb dictionary. For example, the phrase "carried out air strikes against" would be assigned the code 195, whereas "carried out strikes against" be assigned the code 190. To create my own initial verbs dictionary, I combined verb dictionaries from a number of other projects using Tabari for research having to do with conflict in the Middle East, civil war, and/or counterinsurgency operations.

Tabari achieves two goals using these dictionaries. The first goal is to read every lead to determine whether it describes an event (i.e., whether it contains two actors and a verb phrase, as defined by the user-supplied dictionaries, combined in a syntactically meaningful way, as defined by Tabari's text parsing engine). If it finds an event in a given lead, the second goal is to accurately determine who is the source of the action, who is the target, and what the former is doing to the latter, and to transform this into one or more simple codes, in my project taking the form USA #### IRQ and/or IRQ #### USA, where #### is the three or four digit number from one of the sub-categories described above.

As the Tabari coder, my job was to edit the dictionaries to make Tabari read leads more accurately. Accuracy here meant more than one thing: First, it meant minimizing the number of leads describing relevant events that Tabari ignored, while maximizing the number of leads containing no or irrelevant events that it ignored. Second, it meant making sure that verb phrases were properly identified and associated with the correct numeric code. Third, it meant ensuring that Tabari correctly identified all of the relevant actors, and that it did not confuse the source of an action with the target or vice versa.

Creating initial dictionaries was thus just the first step in coding. The next consisted of refining these dictionaries so that they accurately read the specific leads I'd downloaded. To refine dictionaries, I monitored Tabari as it read a few hundred leads, one by one, and constructed event codes from them. When I found an instance of Tabari incorrectly coding (or failing to code)

an event, I edited the dictionaries (usually by adding actors or verb phrases) such that it coded the event correctly. At the end of this process, according to Schrodt, the success rate is about 90 percent, which is about the same as human coding, albeit ultimately much faster.

Once I was reasonably confident of the accuracy of the dictionaries, I ran Tabari on the lead sentence of every AFP newswire submission from April, 2003 to January 2007 containing the word "Iraq" — about 100,000 in total. Tabari transformed this massive text file (28 MB) into event data in only a few minutes.

I then transformed these event data into two time series representing the behavior of each belligerent towards the other over time. To do this, I assigned to each event code a number between -10 and 10, corresponding to a schema developed by Joshua Goldstein and Philip Schrodt, where -10 indicates a high level of hostility and 10 indicates a high level of cooperative behavior from the source towards the target actor. Goldstein based these weightings on the recommendations of a panel of academics who study international politics. Some examples of the weightings are: -10 for a military attack, -9.2 for a seizure of possessions, -6.9 for a threat or an ultimatum, 1.8 for an apology, 7.4 for the extension of economic aid, and 8.3 for the extension of military assistance. I reversed the scale so that higher numbers indicate higher levels of hostility. To create weekly data, I summed numbers assigned to each event code in a given week to create a measure of aggregated, directed levels of hostility. Figures 3.3 and 3.4 are time series data representing US hostility

towards Iraqi insurgents and of insurgent hostility towards the US. There is no visible time trend in either of these series, and augmented Dickey Fuller tests indicate that they are both stationary.

3.4 Empirical Model

The central puzzle of the current project is to identify the determinants of political actors' behavior (in this case, the behavior of insurgents, counterinsurgents, and mass political actors) in an international conflict setting (in this case, the war in Iraq), in order to determine whether these patterns of behavior are consistent with what is suggested by the existent literature. As I suggested in the introduction, two considerations often missing from the empirical analysis of such phenomena are *dynamics* and *simultaneity*. To account for these factors requires an empirical model that allows for the possibility that belligerents' behavior is determined in part or in whole by 1) their own past behavior, 2) the past behavior of the other belligerents, and 3) the past evaluations of mass political actors. Likewise, the model must allow for the possibility that the behavior of mass political actors is determined in part or in whole by 1) their own past behavior and 2) the past behavior of the belligerents. Finally, the model must allow for the possibility that all of these patterns of behavior are simultaneously endogenous — i.e. at once affecting and affected by all of the others. I explain each of these relationships in greater depth in the following chapter.

If k is the order of autoregression, the relationships among the variables

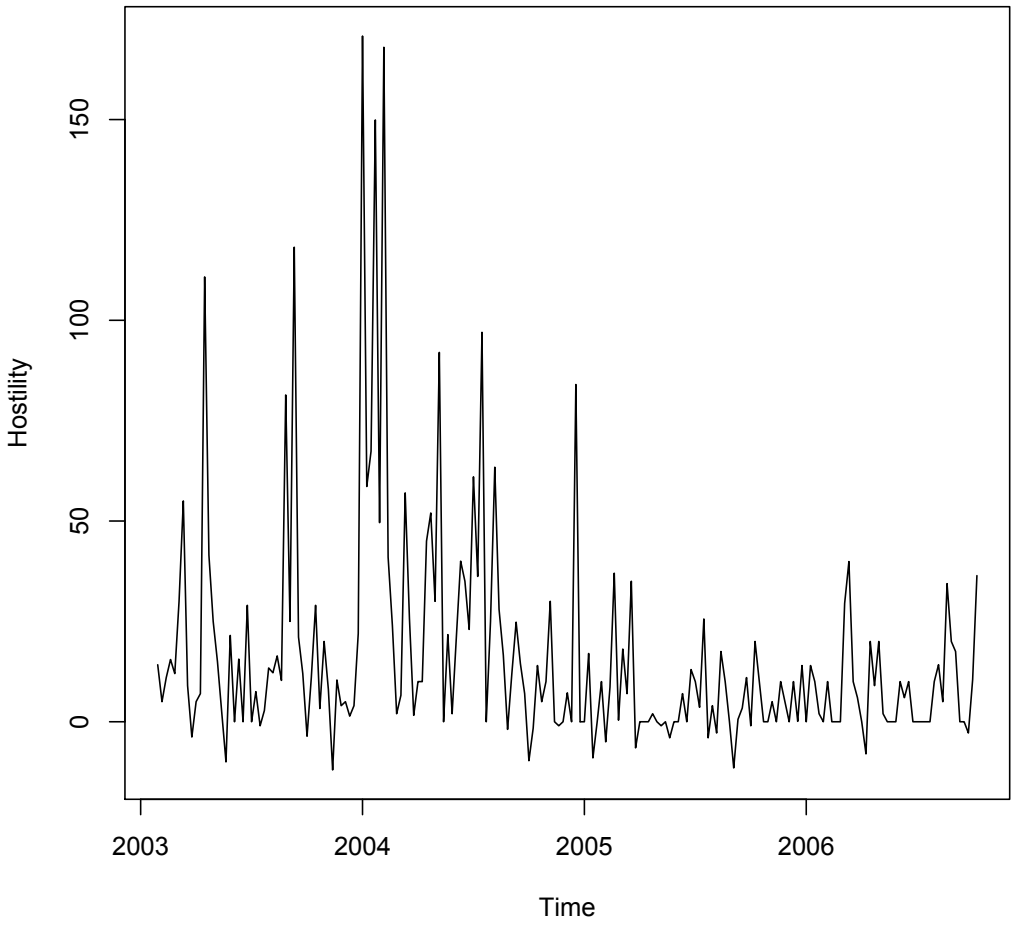


Figure 3.3: Total level of aggression or cooperation, US to Iraq

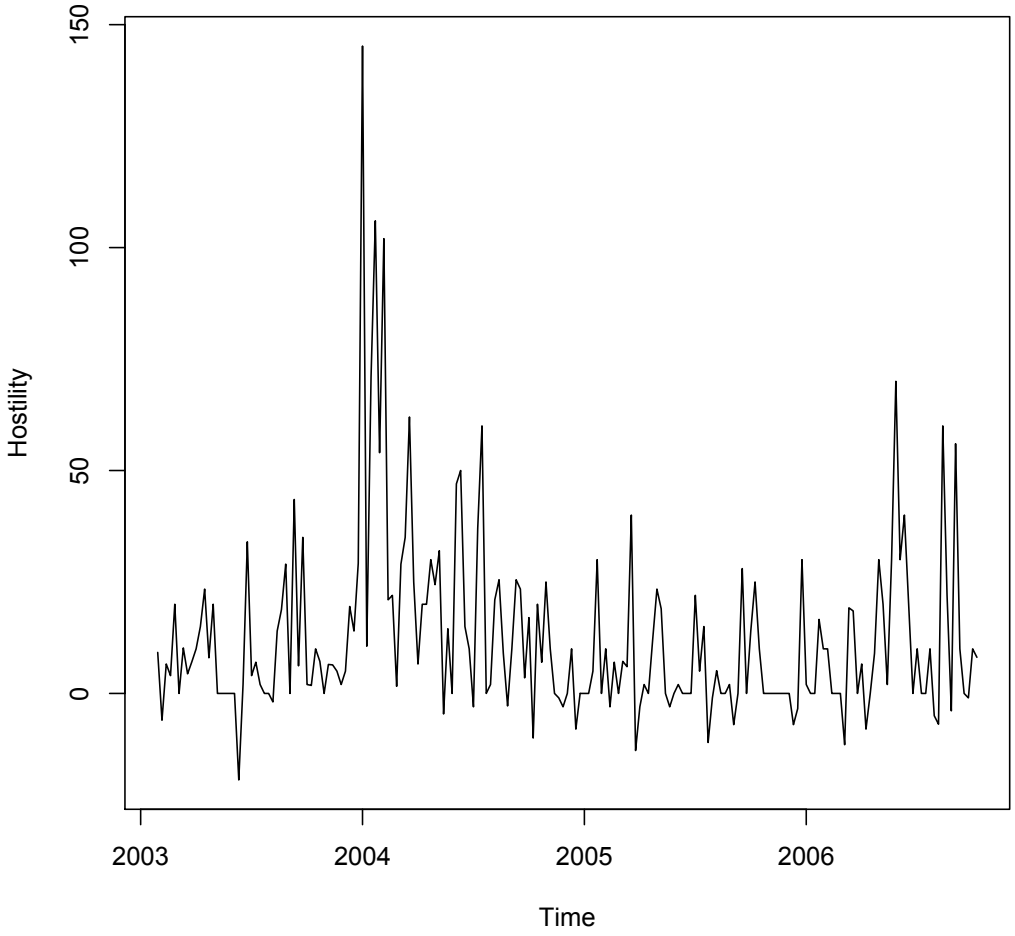


Figure 3.4: Total level of aggression or cooperation, Iraq to US

of interest can be expressed in the following system of linear equations:

$$
\begin{aligned}
Irq2usa &= \alpha_{II} L_K(Irq2usa) + \alpha_{IU} L_K(Usa2irq) + \alpha_{IF} L_K(Fp) + e_I \quad (3.1)\\
Usa2irq &= \alpha_{UI} L_K(Irq2usa) + \alpha_{UU} L_K(Usa2irq) + \alpha_{UF} L_K(Fp) + e_U \quad (3.2)\\
Fpo &= \alpha_{FI} L_K(Irq2usa) + \alpha_{FU} L_K(Usa2irq) + \alpha_{FF} L_K(Fp) + e_F \quad (3.3)
\end{aligned}
$$

Where $Usa2irq$ is the US armed forces' average level of hostility directed toward the Iraqi insurgents during week t, $Irq2usa$ is the Iraqi Insurgents' average level of hostility directed at the US during week t, and Fpo is the change in American public opinion of Bush's handling of foreign policy between week $t-1$ and week t, $L_K(.)$ is an operator that transforms the enclosed variable into a $k \times 1$ vector of lagged variables for $t-1$, $t-2$ up to $t-k$, and the αs are corresponding $1 \times k$ vectors of unknown parameters to be estimated.

For further notational simplicity, these equations can be expressed in matrix form as:

$$\mathbf{Y_t} = \mathbf{A_1 Y_{t-1}} + \mathbf{A_2 Y_{t-2}} + \ldots + \mathbf{A_p Y_{t-p}} + \mathbf{e_t}$$

This representation allows for the estimation of vector autoregressive (VAR) models (Sims, 1980; Freeman et al., 1989). VAR models allow researchers to study the relationships among variables where there is little in the way of theory to guide initial assumptions about *how* those variables might

be related. In particular, which variables are endogenous and exogenous, and each equation's order of autoregression, need not be known in advance of estimation.

3.4.1 Estimation

Because in each equation, the left hand side variable is a function of the right hand side variables at earlier time periods — i.e., they enter the equations as "givens" — there is no autocorrelation, and therefore they can be estimated using ordinary least squares. However, there are some peculiarities in interpreting the results of VAR models, which I will detail below.

These equations are in reduced form. This means the direct and indirect effects of each variable on the other variables, as well as an explicit adjustment for the dampening of these effects over time, are collapsed into the single vector of coefficients A. In principle, it is possible to uncover these structural parameters from a system of reduced form equations by making sufficient restrictions on each of the three equations that the free parameters can all be mathematically identified — i.e., by having an explicit theory about how the variables are related to each other. Unfortunately, under the current circumstances there is no theoretical justification for making such restrictions. But as it turns out, it is possible to learn a great deal about the relationships among these three variables by looking at reduced form coefficients only, as I explain in the next section.

In the absence of a strong theoretical expectation about the value of

p, it is necessary to employ some other criterion for determining lag length. It is standard to use a typical goodness of fit statistic, but special care must be taken in determining what statistic to use. R-squared and log likelihood statistics do not adequately account for the fact that fit can appear to improve as degrees of freedom go down (i.e. as additional parameters are introduced). Akaike's information criterion is a measure of goodness of fit that penalizes the loss of degrees of freedom:

$$AIC = ln\left|\tilde{\Sigma}\right| + \frac{2}{T} \times K$$

where $\tilde{\Sigma}$ is the estimated covariance matrix, T is the number of observations, and K is the number of parameters. The system of equations with the number of lags producing the lowest AIC is the best model according to this criterion. In this case, the criterion selected the model with 5 lags. Thus, my analysis will be based on a VAR(5) model.

3.4.2 Interpretation

Some special care must be taken in evaluating the results of a VAR model. Because of the extreme multicolinearity attending equations containing multiple lags of the same variables, the usual measures of the significance, sign, and magnitude of the effect of each variable on the others are meaningless. Luckily, there are a number of well established approaches for uncovering analogous concepts in VAR models.

Block F-tests or likelihood ratio tests are the primary means of determining the significance of the influence of each variable upon the other. A variant on such tests, the Granger Causality Test, also serves to identify which variables are endogenous, by determining whether the F-tests suggest that the variables are unilaterally dependent, bilaterally dependent, or independent. While not to be confused with proper causality, the endogeneity identified by Granger Causality Tests is a good first cut at determining what are likely to be the primary driving forces in a given data set.

A rough equivalent to measures in the standard regression model of the *sign* (i.e. positivity or negativity) of influence of a given variable on another can be discerned through some form of "innovation accounting" using impulse response functions. Impulse response functions are simulations that charts the response of the system of equations to at shock (usually of one standard deviation) to a variable's residual. These responses can be either positive (suggesting a kind of reciprocity) or negative (suggesting an inverse reciprocity).

Impulse response functions give some idea of both the *magnitude* as well as the sign of the influence of one variable on the others. Another, complimentary measure of this magnitude can be determined by decomposing the forecast error variance of each endogenous variable. This decomposition describes the proportion of future values of a variable that is accounted for by its own and other variables' current values.

I employ all of these techniques to analyze the system of equations

described above in the following chapter.

3.5 Theoretical implications

As I will show in the following chapter, the "audience costs" model is a poor tool for making sense of the insurgency and counterinsurgency in Iraq. The derivation of theoretical implications from this finding requires consideration of where the "audience costs" model breaks down, and why. Certainly "audience costs" *should* matter, especially in modern democracies where voters can remove from office foreign policy makers who make decisions that they do not like. But as it turns out, the influence of domestic audiences is sometimes much less than the "audience costs" literature presupposes. The informational contribution of domestic audiences to the Iraq war is insignificant. There is, furthermore, very little room for differences of interpretation of US domestic politics. This is because the inexorable decline of US public opinion of the war is easily observable for all actors and is therefore essentially common knowledge, and there can be little doubt of the ultimate repercussions for any given administration of this decline. Thus, the extent to which insurgents and counterinsurgents might have different beliefs about the significance of US domestic politics is minimal to nonexistent.

If there is little reason for US counterinsurgents or Iraqi insurgents to pay heed to US domestic audiences, and little room for disagreement between belligerents about the significance of those audiences, then there is little that we can learn about insurgent or counterinsurgent behavior by study-

ing domestic politics in the US. But what of domestic politics in Iraq? The role of domestic audiences in Iraq in determining outcomes for insurgents and counterinsurgents cannot be informational because the nature of domestic audiences' responses to the conflict is never common knowledge. Few if any consistent, reliable and mutually agreed upon measures of these responses exist, for belligerents or for academics. But this suggests that the primary role for Iraqi domestic politics in determining outcomes for insurgents and counterinsurgents could well be in influencing those belligerents' *beliefs*. I explore this possibility in chapter 4.

Chapter 4

Were Bush's critics a fifth column?

4.1 An emboldening effect?

Over the course of the seven years that elapsed between the terrorist attacks of September 11th, 2001, and the departure of President George W. Bush from office, a common claim made by his administration and its supporters was that criticism of his policies undermined US armed forces' ability to wage and win the "war on terror," with Bush suggesting that his domestic critics "emboldened" America's enemies and "dispirited" America's troops and supporters abroad (Bush, 2001). On its face, this claim bears some similarities to the arguments advanced within IR scholarship on the role of "audience costs" in international politics, in particular Kenneth Schultz's work suggesting that democracies' domestic politics contribute directly or indirectly to their actual or perceived resolve (Schultz, 1998, 1999, 2001a). If true, these arguments may have profound implications for how democracies should conduct themselves — both domestically and internationally — in times of international crisis. But whether they are in fact true has not been well established. In what follows, I examine the theoretical bases of the audience costs literature and how it might be brought to bear upon the question of an "emboldening" and or "dispiriting" effect. Then, using vector autoregression analysis of time series data, I

find that there is little evidence to support either such effect, which suggests that domestic audience costs as traditionally understood may have had little bearing on the Iraq war, and perhaps similarly limited explanatory value for insurgencies and counterinsurgencies in general. I conclude by contemplating what, in the absence of audience costs as traditionally understood, might have driven the dynamics of the insurgency and counterinsurgency in Iraq, a question I explore in greater depth in chapter 3.

In the aftermath of the terrorist attacks of September 11th, 2001, public dissent among Americans towards their government's foreign policy was rare. That the American public "rallies around the flag" in times of international crisis is oft-observed (Mueller, 1973, 1994), and this particular rally was matched in intensity only by that following the Japanese attack on Pearl Harbor. According to polling by ABC, within two weeks of President Bush's declaration of a "war on terror" on September 20th, approval of his administration's handling of foreign affairs had soared to 92%, considerably higher than Franklin Roosevelt's peak of 84% (Roper Center, 2007). However, despite the marginality and paucity of American opinions critical of the Bush administration during this period, supporters of its policies nonetheless focused no small amount of energy on responding to and discouraging public articulations of dissent.

Many of the earliest such arguments were reminiscent of — and, indeed, sometimes explicitly invoked — George Orwell's comments on those among his fellow leftists who protested Britain's declaration of war against Hitler's

Germany:

> Pacifism is objectively pro-Fascist. This is elementary common sense. If you hamper the war effort of one side you automatically help that of the other. Nor is there any real way of remaining outside such a war as the present one. In practice, "he that is not with me is against me." (Orwell, 1968).

President Bush used a similar phrase in a speech delivered on September 20th — "either you are with us, or you are with the terrorists" — although unlike Orwell's, Bush's words were directed at the international community rather than at domestic critics (Bush, 2001). But a number of commentators were already making the connection on his behalf. Sunday Times columnist Andrew Sullivan worried about the American "decadent left" forming a "fifth column" (Sullivan, 2001). Explicitly invoking Orwell, Washington Post columnist Michael Kelly wrote

> Organized terrorist groups have attacked America. These groups wish the Americans to not fight. The American pacifists wish the Americans to not fight. If the Americans do not fight, the terrorists will attack America again. And now we know such attacks can kill many thousands of Americans. The American pacifists, therefore, are on the side of future mass murders of Americans. They are objectively pro-terrorist (Kelly, r 26).

Such suggestions were not limited to the comments of pundits alone. Similar language was employed by high-ranking officials in the Bush administration, as well as by Republicans in Congress. In testimony before the Senate Committee on the Judiciary on December 6, 2001, regarding Democrats' concern about the domestic anti-terrorism measures the Bush administration was advocating, then Attorney General John Ashcroft remarked

> [T]o those who scare peace-loving people with phantoms of lost liberty; my message is this: Your tactics only aid terrorists — for they erode our national unity and diminish our resolve. They give ammunition to America's enemies, and pause to America's friends. They encourage people of good will to remain silent in the face of evil (Ashcroft, 2001).[1]

When Democrats continued to voice concerns about the Bush administration's response to the terrorist threat, Republican representative Tom Davis suggested that "divisive comments have the effect of giving aid and comfort to our enemies by allowing them to exploit divisions in our country," and White House communications director Dan Bartlett remarked that the Democrats were doing "exactly what our opponents, our enemies, want us to do" (Keefer, 2002a).

As the war on terror was refocused on the toppling of Saddam Hussein's regime, this rhetorical tactic was extended to include critics of the American

[1]Note, however, that in the same speech Ashcroft also criticized those on the right who claimed that these measures did not go far enough.

invasion and occupation of Iraq. Students for Academic Freedom founder David Horowitz's was perhaps the most extreme rhetorical attack on foreign policy dissenters. Three days before the invasion, he wrote

> the peace movement is not about peace, it is a fifth column communist movement to destroy America and give victory to our totalitarian enemies. Now this Fifth Column is preparing to move into action to attempt to defeat America in its war against Saddam (Horowitz, 2003).

During his 2004 campaign for reelection, George Bush used slightly more gentle language to reprimand his opponent, John Kerry, for criticizing his foreign policy decisions, but nonetheless seemed to be making a similar point:

> You can embolden an enemy by sending a mixed message. You can dispirit the Iraqi people by sending mixed messages. You send the wrong message to our troops by sending mixed messages (Bush, 2001).

What emerges from a review of these and similar comments is a clear pattern of identifying dissent as a form of betrayal with the potential ultimately to undermine national security.[2] These remarks advance, albeit somewhat obscurely, an argument about causality. They suggest that there is a relationship

[2]See Keefer (2002a); Nyhan (2002); Keefer (2002b) for more examples.

between domestic support for an administration at war and the ultimate success of the war effort. But is there any evidence that such a relationship exists?

Much of contemporary international relations scholarship — in particular the literature on audience costs developed by Thomas Schelling (1960), James Fearon (1995), Kenneth Schultz (1998, 1999, 2001a), and others — suggests that it might, and a recent paper by Iyengar and Monten (2008) provides some preliminary evidence that it does. Drawing upon theories of audience costs that tie the ability of belligerents to signal resolve to domestic political processes, Iyengar and Monten develop an empirical model that seems to evince a relationship between public criticism in the United States of the Iraq war and the behavior of America's adversaries in Iraq: Insurgents engage in higher levels of violence during periods when domestic dissent against the war appears greatest. In what follows I will examine the theoretical bases and empirical legitimacy of these claims.

4.2 Signaling resolve

How do members of the Bush administration and its supporters imagine that public dissent, "mixed messages" from the non-incumbent party, or other signs of domestic discontent with US foreign policy affect the behavior of America's adversaries, America's military forces, and/or America's supporters abroad? Answering this question in a manner that generates hypotheses that can be supported or falsified by empirical research requires a great deal of

supposition, as the remarks cataloged above do not offer any explicit theory of causality, and do little to suggest what we might observe if the relationships they describe obtain.[3] What is the precise mechanism that connects public dissent to the emboldening of an enemy or the dispiriting of troops and supporters? What forms of public dissent are likely to result in emboldened enemies or dispirited troops? What do emboldened enemies and dispirited troops look like, and how might they differ from enemies and troops that have not been emboldened and dispirited? Neither Bush nor his supporters ever answered any of these questions directly, but IR scholars have suggested some mechanisms, consistent with the emboldenment hypothesis, that might.

As discussed in the previous chapter, scholars of international relations have often cited Saddam Hussein's statement to U.S. Ambassador April Glaspie in the run-up to Iraq's invasion of Kuwait in 1990 — "Yours is a society that cannot accept ten thousand dead in one battle" (Stein, 1992) — as evidence that Hussein believed democracies (or, at the very least, America's particular form of democracy) are more sensitive to war casualties than other kinds of states, and furthermore that this sensitivity puts them at a strategic disadvantage in militarized disputes (Mueller, 1994; Reiter and Stam, 2002). One can find even more extreme articulations of similar sentiments in the writings of many early– and mid–20th century revolutionaries and insurgents, such as the Vietnamese communist Truong Chinh's assertion that the French

[3]Of course this is also true of a great deal of arguments in the academic IR literature (Wagner, 2007).

people would, given time, so "strongly oppose" the war in Indochina that they would "rise up to overthrow" their own government (Chinh, 2001). If Hussein, Truong, and other potential belligerents, hold such beliefs, then these beliefs could have some effect upon their behavior towards democracies such as the United States and France. Perhaps, then, when the Bush administration and its supporters suggest that disagreement within the United States on questions of foreign policy embolden our enemies and dispirit our troops, and that dissenters are therefore a "fifth column" actively undermining the war effort, what they mean is that our enemies and troops alike hold beliefs such as Hussein's and/or Truong's, that public dissent at home encourages such beliefs, and that we can expect these beliefs to result in a certain pattern of behavior. But what kind of pattern should we then expect to observe, and why?

That America will be particularly sensitive to the costs of war is an assumption central to much of the "audience costs" literature discussed in the previous chapter. This literature holds that the domestic politics of democratic states can, indeed, serve as a rich signal of those states' intentions, capabilities, and resolve. Hussein seems to be primarily concerned with *mass* politics — it is an expansively defined American "society" that cannot bear the war casualties he claims to be prepared to inflict upon the United States. But in much of the "audience costs" literature, it is the behavior of political elites — usually the most prominent public supporters and opponents of the incumbent regime — that communicates the most useful information to potential belligerents. This choice of focus is typically due not only to the assumption that political

elites are better informed than are the masses, but also that elites are capable of cooperation and strategic choice in a way that masses are not (see the discussion in chapter 1 of coordination and collective action for reasons why this is likely to be the case).

Iyengar and Monten's paper is consistent with this elites–focused literature, examining the informational role played by a specific set of political elites — the news media. Dividing Iraq into different regions with varying levels of access to American news media outlets, Iyengar and Monten find that regions with greater access to American media experience greater levels of insurgent violence following the appearance in the media of negative or otherwise "emboldening" statements about the war. As the authors concede, their results may be driven by some unmeasured third factor related to both the levels of insurgent violence and the perceptions of U.S. domestic audiences. As I show below, a third factor does indeed pertain: *U.S. military forces* are responsive to U.S. domestic opinion of the war. Had Iyengar and Monten accounted for the role of changes in the behavior of U.S. forces, their findings might have been quite different. It may well be that it is *only* the U.S. forces that are directly responsive to negative statements about the war in the U.S. media, and that insurgents are responding not to "emboldening" statements in the media, but rather to the changes in the behavior of U.S. forces caused by such statements. I provide evidence suggesting that this might be the case below.

That America's adversaries and soldiers alike might look at the behavior and rhetoric of political elites — in particular, the representatives of both

the incumbent and opposition parties — to guess the future direction of the country's foreign policy would appear to also be the argument implied by the statements made by the Bush administration and its supporters. But to explain why an international counterpart might care about opposing perspectives among political elites in a democracy, when in fact only one party may have any real control over foreign policy decision making at any given time, each of these elite theories of the relationship between domestic and foreign politics relies directly or indirectly on the institutions of mass, electoral politics. The ability of a party to determine foreign policy is dependent upon its members gaining and maintaining power, which is in turn dependent upon their policies remaining popular with a plurality of voters. Thus, when public opinion of foreign policy is high and/or increasing, this would be a fairly clear indication that the policy is likely to continue. Likewise, when public opinion of foreign policy is low and/or decreasing, this would be a clear indication that the policy is likely to soon end.

The best measure of America's resolve, then, is not the content of the opposition's critique of the incumbent party's foreign policies, but rather the general public's opinion of those policies. What information, then, can the behavior of political elites communicate that is not already communicated by mass opinion? As mentioned above, Schultz and other contributors to the "audience costs" literature have suggested that these elites are better informed and more capable of thinking and acting strategically than are the general public. But regardless of whether they are better informed or better able to

make strategically informed choices, if they are unable to either *anticipate* or *influence* mass opinion in a way that other actors cannot, it is not clear how they could be expected make a unique informational contribution to a conflict setting. In fact, the evidence that they can do either in a way that is relevant to foreign policy decision–making is far from compelling. John Mueller and many others have shown, for example, that American opinion of involvement in foreign wars is more or less insensitive to the anti–war sentiments of political and media elites (Mueller, 1973). Instead, Mueller suggests, public opinion decreases predictably, consistently, and ineluctably as casualties mount.[4]

Despite these caveats, there is significant evidence that at least one group of political elites believed — or at the very least hoped — it could influence public opinion of the war: the Bush administration itself. While they were advising the Bush administration, Peter Feaver, Christopher Gelpi, and Jason Reifler have provided some evidence for a modification of Mueller's explanation of the relationship between public opinion and war (Gelpi et al., 2007). While they acknowledge that it is generally true that public opinion declines as casualties increase, they suggest that this relationship does not hold when the public can be convinced that the war is likely to succeed. Unsurprisingly, after Feaver et. al. began advising Bush, his speeches increasingly focused on evidence of success in the prosecution of the war in Iraq (Shane, 2005). While Bush was ultimately unsuccessful in making this case, as I argue below, some

[4]More precisely, Mueller suggests that public opinion decreases as a logarithmic function of war casualties.

of the behavior of US forces in Iraq can be best explained as a response to, and an attempt to reverse, the then already apparent downward trend in public opinion.

Regardless of whether political and media elites have the influence suggested by the "audience costs" literature, much of the direction of foreign policy is determined by the general public. Thus, even if America's soldiers and adversaries can and do look to the proclamations of political and media elites for some of the more subtle cues about resolve and the future direction of foreign policy, they should still also pay some attention to the clearest and most direct signal of such — public opinion. If public opinion regarding a president's foreign policy goes down, this would appear to be the clearest indicator of weakening resolve, and a generally increased willingness of the public to try a different course of action —and to replace the current regime with one willing to pursue such a course. If it is possible to uncover a relationship between public opinion of foreign policy and the behavior of US soldiers and foreign adversaries — that US soldiers or foreign adversaries pay attention to public opinion and respond to it — then this might well imply an emboldening or dispiriting effect. But if there is no evidence that foreign adversaries or US troops pay attention and respond to such a simple and obvious measure of American foreign policy resolve as public opinion, then it seems unlikely that they would consider any of the more subtle measures that are implied by the elite theories outlined in chapter 1.

4.2.1 Bystanders and followers, accountability and credibility

Bush's comments suggest not only that the "mixed messages" of domestic politics can "embolden" the enemy, but that they might also have a correspondingly "dispiriting" effect on US soldiers. Thus, he seems to be implying an overall effect of domestic politics on all belligerents' *levels of aggression* or *hostility*. As it turns out, this is the primary variable of interest in the action-response models common in the empirical literature on reciprocity. Building on the empirical and theoretical models of Richardson (1960) and Axelrod (1981), this literature has shown that the guiding principle in bilateral and trilateral relations among potential belligerents is often one of "tit-for-tat," allowing for mutually beneficial, cooperative behavior even where states are predisposed to distrust one another (Goldstein and Freeman, 1990). In the following sections, I will explain how these models can be adapted to test the hypotheses suggested by the statements of Bush and his supporters.

Brandt, Colaresi, and Freeman have already suggested a number of ways to modify models of reciprocity to account for the possible influence of public opinion on the relations among international rivals (Brandt and Colaresi, 2008). They suggest four possible models of the influence of domestic politics on belligerents' behavior, each nested in the last. The first, which takes as its starting point suggestions by Lippman (1922), Almond (1965), Morgenthau (1967), and Rosato (2003) that the public actually tend to be quite uninformed about foreign policy , is the "bystander" model, where the public pays little attention to foreign policy. The public here largely ignores what its

own and other foreign policy leaders are doing, and therefore foreign policy leaders ignore the public in turn. From the perspective of international politics, then, the bystander model is roughly equivalent to those action-response models that do not allow for a possible influence of public opinion foreign policy decision-making.

The second, inspired by observations of the "rally-round-the-flag" effect as well as diversionary theories of war (see discussion in chapter 1 above), is the "follower" model, where the public, while informed about international political phenomena, gets its information exclusively from its own foreign policy elites. Brandt, et al allow for two possible ways that the public might be followers, one where they tend to approve of any choice made by its foreign policy leaders, and the other where they tend to approve of initial decisions but disapprove of any change of course that foreign policy-makers might take thereafter. In either case, foreign policy leaders once again have little reason to pay any heed to public opinion — as a predictable response to decisions those leaders have already made, it contains no new information. Once again, from the perspective of international politics at least, this model is roughly equivalent to one where public opinion is left out entirely. But unlike the previous model, here international politics can be used to explain domestic politics.

Given that the public could either approve or disapprove of decisions made by their own or other foreign policy leaders, and that these opinions can have considerable impact on the future political careers of these leaders, two other models suggest themselves. The first is what Brandt et al call the

"accountability" model. Rooted in the Madisonian tradition of understanding democratic institutions as primarily a means of constraining leaders the accountability model suggests that domestic political actors observe their foreign policy decisions and remove them from office when they do not like the decisions that they have made. Because foreign policymakers can anticipate the ejection from office that a dip in public opinion may foreshadow, they should be responsive to such fluctuations, changing their future behavior when the response to past behavior has been negative.

If foreign policymakers are held accountable for their actions by their voters, and therefore observe and respond to public opinion, then it may well follow that their international adversaries can find hints about their future behavior by paying attention to the consequent fluctuations in public opinion. This "credibility" model would appear to be more or less identical to the audience costs models suggested by Schelling, Fearon, and Schultz, only applied to bargaining while fighting rather than crisis bargaining.

4.2.2 Can an uninformed public inform?

These models are far from exhaustive in their treatment of plausible relationships between domestic and international politics. In particular, Brandt, et al's models do not acknowledge that an ignorant and/or easily manipulable public — as in their "follower" and "bystander" models — can nonetheless have enormous influence on the direction of foreign policy — as in their "accountability" and "credibility" models. In modern, liberal democracies, the

right to vote is not reserved for the most well informed and autonomous citizens; the uninformed and compliant can also vote. Thus, even the opinion of a voting public that evinces little or no awareness of political realities may ultimately matter a great deal for foreign policy leaders — uninformed or deluded voters can reelect or remove them from office just as easily as informed voters would.

This relationship can lead to rather perverse outcomes, such as in the principal–agent models analyzed by Downs and Rocke (1994). If, for example, an uninformed public's metrics for evaluating their leadership are poor — such that they consider only the *success* of a leader's choice rather than whether the choice was *ex ante* a good one — then they may punish good but unlucky leaders while rewarding bad but lucky ones. This perverse pattern of behavior from the public leads to even more perverse patterns of behavior from foreign policy leaders, who, knowing they are likely to be kicked out of office if they fail, follow bad or unlucky decisions with even worse decisions in an attempt to "gamble for resurrection." As I suggest below, this is one way of looking at the majority of US counterinsurgency in Iraq — as an extended and ultimately unsuccessful gamble for resurrection.

4.3 Analysis

Bush's statements that mixed messages "embolden an enemy" imply that we should expect to see waning public support result in higher levels of hostility from Iraqi insurgents, and therefore that increasing support should

result in lower levels of hostility. His statement about "dispiriting" the troops similarly implies that waning public support should result in lower levels of hostility from American forces: Demoralized by a lack of support at home, US troops would not fight as hard.

In other words, the following relationships should obtain:

- The vector of parameters represented by α_{IF} in equation 3.1 should be jointly significant.
- Impulse response functions[5] based on the system equations 3.1, 3.2, and 3.3 should indicate a negative effect of Fpo on $Irq2Usa$.
- The vector of parameters represented by α_{UF} in equation 3.2 should be jointly significant.
- Impulse response functions based on the system equations 3.1, 3.2, and 3.3 should indicate a positive effect of Fpo on $Usa2Irq$.

Given the results from the literature on reciprocity discussed above, we should also expect the following:

- The vector of parameters represented by α_{IU} in equation 3.1 should be jointly significant.

[5]See below.

- Impulse response functions based on the system equations 3.1, 3.2, and 3.3 should indicate a positive effect of $Usa2Irq$ on $Irq2Usa$.

- The vector of parameters represented by α_{UI} in equation 3.2 should be jointly significant.

- Impulse response functions based on the system equations 3.1, 3.2, and 3.3 should indicate a positive effect of $Irq2Usa$ on $Usa2Irq$.

And an open question is whether the public consists primarily of "bystanders" or "followers." If the latter, we should expect either the vector of parameters represented by α_{FU} to be jointly significant or the vector of parameters represented by α_{FI} to be jointly significant, or both. If neither vector is jointly significant, this suggests a "bystander" model, but it does not rule out the possibility of "gambling for resurrection" or other principle–agent problems.

As noted in chapter 3, because VAR models contain multiple lagged versions of the same variables, the consequent multicollinearity makes any individually estimated coefficient and its accompanying standard error effectively meaningless. The resulting difficulty in establishing anything analogous to the standard regression model's usual measures of magnitude and statistical significance of influence is the primary stumbling block for the employment of VAR models. A number of alternative approaches for interpreting results exist: Granger causality tests, decomposition of the forecast error variance, and graphing impulse response functions.

Granger's causality test (Granger, 1969; Sims, 1972) is essentially a series of bivariate/block F-tests, to determine if the coefficients associated with the regression of one variable, and its lags, on another are jointly significantly greater than zero. Granger causality tests are a good first cut at uncovering the relationships among variables, but because they are essentially bivariate, they do not properly account for multiple endogeneity in all of its complexity. Thus, they are not particularly instructive in the present context. Nonetheless, I have included in table 4.1 the results of Granger causality tests for the three variables of interest. The Granger tests show that changes in US hostility towards Iraqi insurgents appear to have a direct effect on changes in Iraqi insurgents' hostility towards the US and vice versa. That is, there would appear to be some form of reciprocity between both sets of hostiles. They also show that changes in public opinion may have some (small) direct effect on changes in US hostility towards Iraqi insurgents. Importantly, neither US hostility towards Iraqi insurgents nor Iraqi insurgents' hostility towards the US seem to have any effect upon public opinion. And most important to the emboldening effect hypothesis, there is no evidence that changes in public opinion Granger cause changes in Iraqi forces' level of hostility. However, consistent with the dispiriting effect hypothesis, there is some evidence that changes to public opinion do Granger cause changes in US forces' level of hostility, though it is unclear in what direction this effect is.

Decomposition of the forecast error variance allows for a measure of the percentage of the variance in the outcome variable at time $t + s$ that is

	F-statistic	p-value
irq2usa → usa2irq	4.633	0.003
fpapp → usa2irq	2.094	0.101
usa2irq → irq2usa	4.161	0.006
fpapp → irq2usa	0.410	0.746
usa2irq → fpapp	0.226	0.8782
irq2usa → fpapp	0.247	0.863

Table 4.1: Results of Granger test

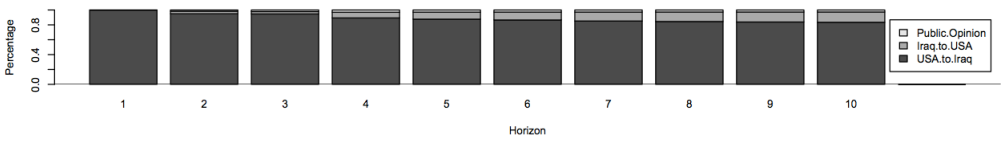


Figure 4.1: Decomposition of Forecast Error Variance for US Hostility

attributable to an orthogonal innovation at time t. Figures 4.1, 4.2, and 4.3 show that by $t+10$, still over 98 percent of the variance in public opinion is due to itself, almost completely engulfing effects from the other variables. While the Granger test indicates that there is some impact of changes in public opinion on the behavior of US forces, forecast error variance decomposition shows that this effect is actually quite small. By time $t + 10$, about 3 percent of the variance in US forces' hostility towards Iraqi insurgents is due to the initial shock. By this time, only .6 percent of the variance of Iraqi forces'

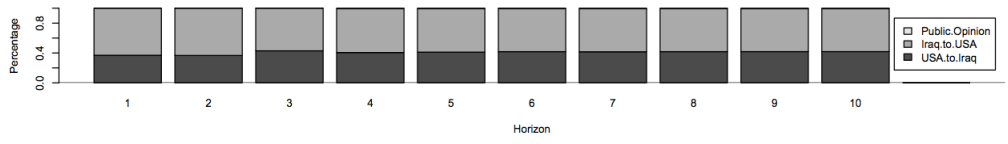


Figure 4.2: Decomposition of Forecast Error Variance for Iraq Hostility

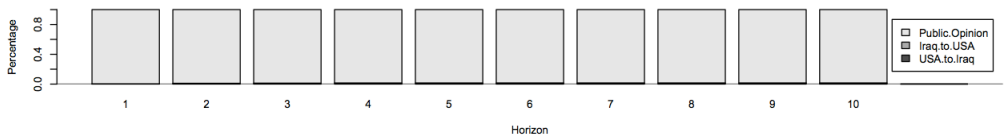
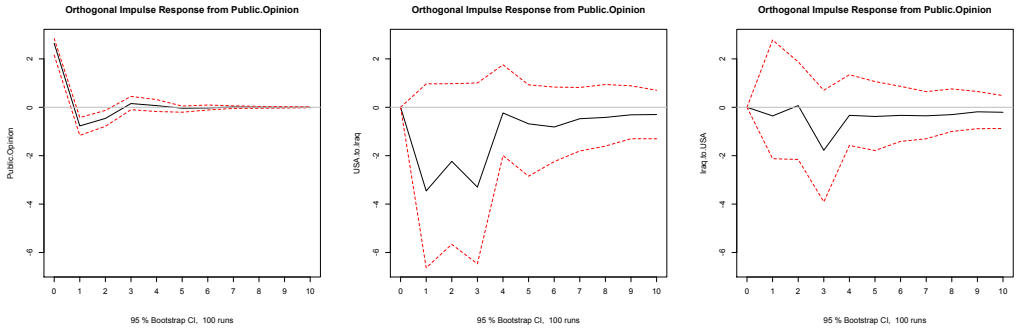


Figure 4.3: Decomposition of Forecast Error Variance for Public Opinion

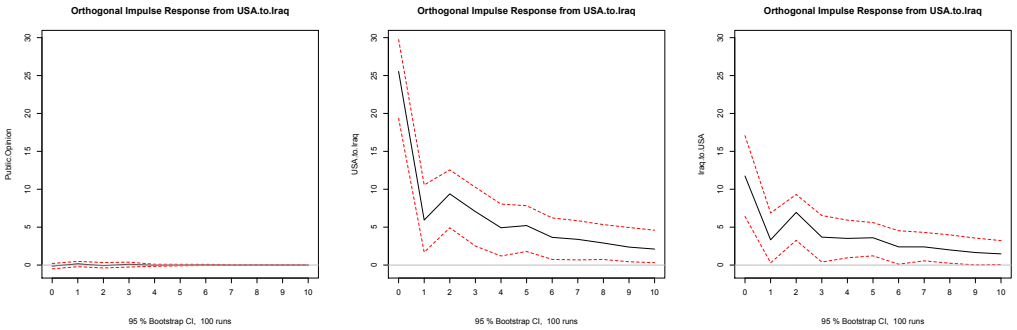


(a) Public Opinion (b) US Hostility (c) Iraq Hostility

Figure 4.4: Impulse Response Functions from Public Opinion

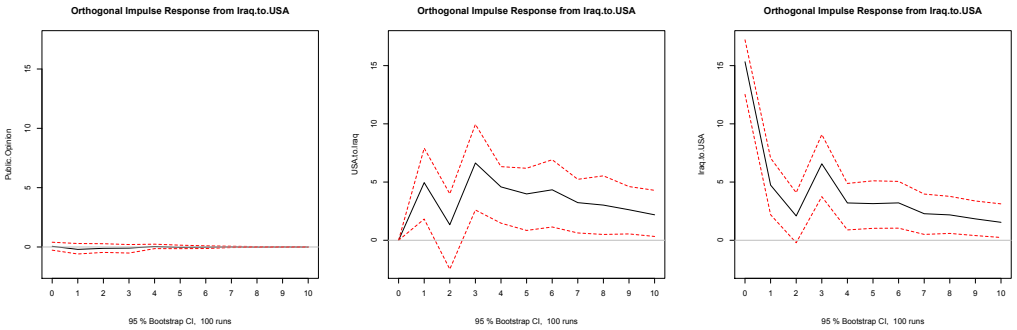
hostility towards US forces is due to changes in public opinion.

Finally, impulse response functions help describe the precise chain reactions implied by the system of equations. Figures 4.4, 4.5, and 4.6 show how a shock to one variable reverberates in another. While some have suggested that the attitude of insurgents towards counterinsurgents is often characterized by "aggressive opportunism" — where insurgents' hostility towards the counterinsurgent increases (decreases) in response to decreases (increases) in counterinsurgent hostility — in the Iraqi context, counterinsurgents and insurgents appear to have a reciprocally hostile relationship similar to those often observed between major powers. That is, as counterinsurgents' hostility increases (decreases), insurgents' hostility increases (decreases) and vice versa.



(a) Public Opinion (b) US Hostility (c) Iraq Hostility

Figure 4.5: Impulse Response Functions from US Hostility



(a) Public Opinion (b) US Hostility (c) Iraq Hostility

Figure 4.6: Impulse Response Functions from Iraq Hostility

But the most notable significant relationship is that between public opinion and the hostility of US forces: decreases (increases) in public opinion lead to an increase (decrease) in the hostility of US forces. Because the relationship between US and Iraqi forces is one of reciprocity, this in turn leads to an increase in the hostility of Iraqi forces. Taken together, these relationships do seem to suggest an effect of public opinion on the behavior of Iraqi forces. The relationship does not, however, seem consistent with the emboldening effect conjecture, because what the Iraqi forces are directly responding to is a change not in public opinion, but in the behavior of US forces caused by changes in public opinion. The results also directly contradict the suggestion of a dispiriting effect: The evidence seems to suggest that US forces become more aggressive, not less, as public opinion of the war goes down.

Clearly the most notable feature of these data is not what they suggest about the behavior of Iraqi insurgents — on the contrary, Iraqi insurgents appear to rather predictably reciprocate the behavior of US counterinsurgents — but what they suggest about the behavior of US counterinsurgents. Contrary to Bush's assertions, it is US counterinsurgents who appear to be "emboldened" by the mounting criticism of US war efforts evinced by flagging public opinion. Why would this be the case? As noted above, Peter Feaver was hired as an advisor by the Bush administration in June of 2005, and perhaps not coincidentally was at the same time in the process of co–authoring a number of articles suggesting that the best way to win public support of a war is to convince the public that the war is a success. One plausible explanation for the

relationship between public opinion and US counterinsurgent activity, then, is that there is an electoral logic driving the latter: Sensitive to the extent to which the continuation of its policy in Iraq would be decided by future elections, the Bush administration responded to dips in public opinion by ramping up its efforts in Iraq to better project an image of success at home.

An even stronger, and for that matter more directly testable, assertion would be that this is not only evidence of the aforementioned electoral logic, but also — contrary to the concerns expressed by Nye and Walt discussed in chapter 1 — of the influence of academia on politics. At the time of his appointment, a number of news analysis pieces suggested the existence of such an influence, implying that the Bush administration's approach to the war in Iraq had altered significantly in the latter half of 2005, and that the change was directly attributable to Feaver's work (Shane, 2005; Baker and Balz, 2005). But is there evidence of a significant change in the behavior of US counterinsurgents after Feaver's appointment to the Bush administration?

A Chow test is one way to determine whether a structural break occurs in the data at about the time of Feaver's appointment. A Chow test is essentially a simple F–statistic based on the residuals from three OLS estimates, in this particular instance of the mean of the time series representing US hostility towards Iraqi insurgents:

$$usa2irq_t = \alpha + u_{1t}, t = 1, 2, \ldots, n_1 \tag{4.1}$$

for the period before the proposed structural break,

$$usa2irq_t = \beta + u_{2t}, t = n_1, n_1 + 1, \ldots, N \tag{4.2}$$

for the period after the proposed structural break, and

$$usa2irq_t = \lambda + u_t, t = 1, 2, \ldots, N \tag{4.3}$$

for the whole time series.

It is then only necessary to calculate

$$F = \frac{(RSS_3 - RSS_1 - RSS_2)/2}{(RSS_1 + RSS_2)/(N-2)} \tag{4.4}$$

where RSS_1, RSS_2, and RSS_3 are the residual sum of squares from equations 4.1, 4.2 and 4.3, respectively. If F is sufficiently large, we can reject the null hypothesis that the means before and after the supposed break — i.e. if F is sufficiently large, then there is a break (Gujarati, 2002).

The F–statistic for a break at mid-June of 2005 is 13.26, corresponding to a p-value of 0.0003. While this is strong evidence in support of a structural break, the Chow test is a very blunt instrument, and does not rule out the possibility that the true break might have occurred at an earlier or a later date — the null hypothesis here is "no structural break" not "no structural break at a different date." For example, the F–statistic for a break at mid-May of 2005 is 16.78, and for mid-April it is 20.83 (and the corresponding

p values go down). In other words, the evidence of a structural break at an earlier date seems even stronger. Intuitively, a breakpoint with a larger corresponding F–statistic would seem to be a better candidate for the moment a structural change occurred. Zeileis et al. (2003) suggest that this intuition is well–founded, and that one should therefore seek an optimal breakpoint, defined as the point where the corresponding F–statistic is largest. Calculating F–statistics for all possible breakpoints in the time series indicates that, in fact, the major change in the behavior of US counterinsurgents occurred in the summer of 2004 ($F = 29.1553$, $p = 1.969 \times 10^{-07}$), a full year before Feaver's appointment. Thus, if there is an electoral logic behind the US approach to counterinsurgency in Iraq, Feaver's appointment is likely a symptom of that logic, not its cause.

4.4 Conclusion

The argument that a nation's success in a conflict depends to some extent on its ability to project an image of unified determination, and that liberal democracies might face difficulties in this regard that other kinds of states do not, is hardly unique to the proponents of the war in Iraq. It was in fact already a well-worn truism by the time of Orwell's voicing of similar concerns a half-century earlier. Concerns about such emboldening and dispiriting effects are not unjustified. But before we have a discussion about the implications of this insight, we first must determine whether it is, in fact, true.

It may well be that in some instances public dissent at home can have

an adverse effect on the morale of American troops abroad, and that decreased morale might hinder troops' ability to fight. This seems to be the meaning of the second part of Bush's reprimand of John Kerry for criticizing his policies in Iraq, that sending a "mixed message" to the troops may dispirit them and thus undermine the war effort. As it turns out, my analysis of the data suggests that, if anything, the opposite is true in Iraq. There appears to be a *negative* relationship between public opinion and the level of hostility US forces display towards their opponents. That is, when public opinion goes down, the willingness of US troops to inflict losses upon their adversaries in Iraq apparently goes up.

It seems unlikely, however, that this would be evidence that decreases in public opinion correspond directly to an increase in American troops' *morale*. A more plausible explanation for this relationship is that it reflects the influence not of public opinion on the troops themselves, but rather of public opinion on the decisions of the foreign policy leaders that command them. As discussed above, Feaver and Gelpi argue that the public will support a war if they believe it is likely to be successful — a conclusion that the Bush administration appears to have shared. Given the electoral logic of American foreign policy making — i.e., that in order to make foreign policy decisions, one needs to gain and maintain political office — the relationship uncovered by the above data analysis is likely evidence that the Bush administration sought to calibrate policy in Iraq so as to influence public opinion of that policy at home. Prior to the introduction and widespread adoption of the "population–

centric" tactics associated with the "surge" (see chapter 5), this would have meant increasing hostility against insurgent forces in response to decreases in public opinion, so as to project an image of battlefield success to observers at home. In other words, this is behavior that is more or less consistent with what Downs and Rocke call "gambling for resurrection."

While the data thus suggest that US foreign policy makers paid attention and responded to fluctuations in public opinion, and anecdotal evidence suggests that they did so with an eye towards influencing it, the relationship does not appear to have been reciprocal. That is, despite their best efforts, US foreign policy makers were not able to change the trajectory of public opinion of the war in Iraq. In other words, the gamble was ultimately unsuccessful.

The data further indicate that Iraqi insurgents were not paying attention and responding to American public opinion of the war at all — at least, not to its short–run fluctuations. What Iyengar and Monten see as evidence of such a relationship is likely indication only of an indirect effect of public opinion on the behavior of Iraqi insurgents *through* US forces. That is, contrary to the assertions of the Bush administration and its supporters, it is only the US forces that responded (or appeared to respond) to fluctuations in public opinion, and the Iraqi insurgents merely responded to the US forces' response.

Despite these findings, it is still quite likely that Iraqi insurgents paid (and perhaps still pay) *some* attention to American public opinion of the war. I suggested above that the very obvious, long–run, downward trend in public opinion had no analogue in the time series of belligerents' behavior. But this

doesn't mean that no such trend in belligerents' behavior existed, only that the data I collected did not measure it. While I cannot provide evidence, I would suggest that the trend in the behavior of Iraqi insurgents corresponding to the downward trend in US public opinion is their increasing confidence that they can win the war against US counterinsurgent forces. Put another way, *the fact that they keep fighting* is the trend corresponding to the downward trend in US public opinion.

In the earlier phases of an insurgency, when the conflict is at its most asymmetrical, conventional military means are not available to insurgents. During these phases, it is therefore necessary for insurgents to adopt what contemporary theorists of counterinsurgency call "population–centric" tactics (see chapter 5 for a more in–depth discussion of what this means): Their focus is less on direct engagement with the counterinsurgent and is instead on winning over the civilian population of the territory they seek to control and, more important to the present discussion, *turning the counterinsurgent's own domestic base of support against her* (as suggested by the discussion of Hussein above). When US public opinion continued to go down, this suggested that the latter part of this strategy, at least, was successful. Thus the trend in insurgent behavior that corresponds to the downward trend in US public opinion is the insurgents increasing belief that the war is worth fighting and, ultimately, winnable.

While this explanation is plausible, it is far from satisfying. It suggests that dynamic, multiply endogenous processes, the search for which motivated

the current project, are nowhere to be found in the current conflict in Iraq — except, trivially, insofar as the insurgents and counterinsurgents evince the same patters of reciprocity that have been observed among rivalrous states (Goldstein and Freeman, 1990). But what about domestic political actors? In the following chapter I attempt to find another role for them in this conflict.

Chapter 5

A Random Walk in the Moonlight: Divergence and Convergence of Beliefs in Insurgency and Counterinsurgency

In the following chapter, I elaborate a model of bargaining and learning while fighting that allows for disagreement among belligerents about how best to measure success on the battlefield. This disagreement can cause belligerents to initially become more, rather than less, mutually optimistic while fighting. Thus, in this model, it is possible for belligerents' expectations about winning to diverge in the short run, before ultimately converging in the long run. This is a fundamental departure from the majority of the literature on mutual optimism, which The model can be used to explain why, in large-N studies of war, the likelihood of war termination initially decreases before increasing (Ramsay, 2008). It also can be used to understand the dynamics of individual wars. In particular, it helps explain the conditions that led to insurgent and counterinsurgent violence during the US occupation of Iraq escalating and then peaking in 2007.

5.1 Motivation

What makes insurgency and counterinsurgency different from other forms of warfare? What is *irregular* about irregular war? The social science literature on insurgency and counterinsurgency evinces widespread agreement that insurgency and counterinsurgency should be considered categories apart from other kinds of war, and equally widespread disagreement about what makes them special. Are the underlying causes of insurgency and counterinsurgency essentially distinct from the underlying causes of conventional war? Are the goals of insurgents and counterinsurgents different from the goals of belligerents in conventional war? Are the means of fighting available to insurgents and counterinsurgents different from the means of fighting available to belligerents in conventional war? If any of these differences do exist, how do they affect the belligerents' choice of tactics? Are the factors contributing to the termination of insurgencies and counterinsurgencies different from those of conventional wars? All of these questions have contestable and regularly contested answers.

These are not, it should be mentioned, purely academic debates. The difficulties that social scientists have in answering these questions are mirrored by the difficulties belligerents face in formulating and implementing consistently successful approaches to the practice of insurgency and counterinsurgency. The similarities between the disagreements of the practitioners and the disagreements of the theorists of insurgency and counterinsurgency should not be surprising. Many of the most prominent theories of insurgency and coun-

terinsurgency come from its practitioners – from Mao Zedong and Che Guevera to David Galula and C.E. Callwell to Jonathan Nagl, Steven Metz, and David Kilcullen. Knowing this, we might assume that the beliefs of the practitioners of insurgency and counterinsurgency about the nature of the conflict in which they were engaged affected the decisions they made while fighting, and likewise that their beliefs were affected by the outcomes of their decisions. Thus the similarities between differences of opinion in theory and differences of opinion in practice are instructive, as they suggest that we might be able to better understand insurgency and counterinsurgency not by attempting a definitive refutation or proof of any of these arguments, but rather by understanding how these disagreements occur in the first place, and how the disagreements themselves might affect a war's outcomes. If the practitioners of insurgency and counterinsurgency are just as confused about the nature of the war they are fighting as are those of us who are trying merely to understand it, then it may well be that this very confusion is theoretically and substantively meaningful.

It is my contention that the confusion evinced by many of these arguments is in fact an inherent and unavoidable attribute of irregular warfare, as it stems from questions that cannot be definitively answered without actually fighting an irregular war to its conclusion. This is because whether, how, by whom, and why insurgencies and counterinsurgencies are fought and won will vary from case to case, and as I will argue many of the factors that determine outcomes are not always easily or immediately observable — not even by the belligerents themselves, and even well after fighting commences. If they are

unobservable, then belligerents must make guesses about them, and may well disagree in how they guess.

Such disagreement can affect the war's outcome in at least two ways: First, it can cause belligerents to overestimate their likelihoods of winning, which may lead them to make excessive and disagreeable settlement demands of each other, thus prolonging the conflict. This is an insight common among contributions to the literature on conflict bargaining that focus on "mutual optimism" as an explanation of war. The "silver lining" in many of these analyses, to the extent that there is one, is the insight that fighting a war provides opportunities for adversaries to learn previously unknown or incommunicable information about their chances of winning. And because the outcomes of a war can be observed by all parties, war can potentially bring their beliefs and expectations into alignment. Thus, mutual optimism can lead adversaries to war, and war can then help adversaries overcome mutual optimism.

But the potential for such *ex ante* disagreement raises the possibility of a second kind of disagreement, which has not been addressed in the formal literature on conflict bargaining. If belligerents disagree about the determinants of success, this very disagreement can lead them to interpret differently even those events that they do both observe. Therefore, they can disagree not only on the interpretation of events that occur before the war begins, but also on the interpretation of events that take place during the war itself. As I will show, these *ex post* disagreements may in the short run undermine the process through which war can otherwise bring their beliefs and expectations

into alignment.

In the following chapter I propose an approach to thinking about insurgency and counterinsurgency that allows for the possibility of both such differences of opinion — *ex ante* and *ex post* — and explains how and under what circumstances they can affect the outcomes of irregular warfare.

5.1.1 Two views of insurgency and counterinsurgency

The majority of contemporary disagreements over the practice of insurgency and counterinsurgency hinges on whether and how to utilize civilian populations. Following David Kilcullen, I divide approaches to thinking about insurgency and counterinsurgency into two categories: "enemy centric" and "population centric" (Kilcullen, 2009).[1] I will elaborate below how these approaches differ.

5.1.1.1 The "population centric" approach

According to much of contemporary counterinsurgency theory, what makes insurgency and counterinsurgency different from conventional warfare is the role played by civilians in determining ultimate success. In a formulation that is usually attributed to Lyndon Johnson, but that has since moved into the realm of cliche, it is the "hearts and minds" of the local population that

[1]Note that Kilcullen uses these terms to describe differing approaches to counterinsurgency only. But as the discussion below makes clear, insurgents are faced with a very similar decision about the utilization of civilian populations. The key difference is that in the early phases of insurgency, an "enemy centric" approach is a less viable option for the insurgent, due to asymmetries in conventional war–making capabilities.

the insurgents or counterinsurgents must win before a victor can be declared. This insight goes back at least as far as that insurgency now known as the American Revolutionary War, as evinced by British General James Robertson's comment: "I never had an idea of subduing the Americans. I meant to assist the good Americans to subdue the bad" (Mackesy, 1993).[2] The notion that a counterinsurgent can and should divide a population into "good" civilians and loyalists and "bad" revolutionaries implies that there is some value in doing so. But what might that be?

The French theorist David Galula provides the clearest articulation of the underlying logic of a population-centric approach. Echoing Rousseau's discussion of the consensual basis of political rule, Galula defines the central problem for both insurgents and counterinsurgents as follows:

> Afflicted with his congenital weakness, the insurgent would be foolish if he mustered whatever forces were available to him and attacked his opponent in a conventional fashion, taking as his objective the destruction of the enemy's forces and the conquest of the territory. Logic forces him instead to carry the fight to a different ground where he has a better chance to balance the physical odds against him. The population represents this new ground. If the insurgent manages to dissociate the population from the counterinsurgent, to control it physically, to get its active support, he will win the war because, in the final analysis, the exercise of political

[2]The comparison between the British approach to fighting American revolutionaries and the American approach to fighting Iraqi insurgents is due to Ricks (Ricks, 2008).

> power depends on the tacit or explicit agreement of the population or, at worst, on its submissiveness (Galula, 2006).

Galula, examining in particular the writings of Mao, also handily sketches out how this approach is put into practice. The key differences between insurgency and conventional warfare manifest themselves in the early stages of the war. These involve first the construction of a strong, cohesive insurgent elite, and then the steady accumulation of support for the insurgency from other groups as well as the unorganized masses. During this period of accumulating mass support, the insurgents' only real advantage is their flexibility and ease of movement. Thus, they tend to act in the manner described by Galula above, avoiding conventional, direct confrontation with the counterinsurgents, and instead focusing on ambush, sabotage, and brief engagements with help from supportive members of the civilian population, some of whom may eventually become insurgents themselves — what Kilcullen calls "accidental guerillas" (Kilcullen, 2009). Thus, in its early phases, insurgency is essentially population centric; a successful insurgency cannot be otherwise. It is only after civilian support has reached a critical point that insurgents will have the resources necessary to fight according to the strategies and tactics of conventional war, and thus become enemy centric (see below). In fact, once this point has been reached, the population centric tactics of insurgency are increasingly less viable, given that insurgent forces will have become as large and inflexible as a regular army (Galula, 2006).

If an insurgency can only be won with the support of the population, it

follows, then, that one approach to successful counterinsurgency might be to prevent the insurgents from winning over the population to their side. From a counterinsurgency perspective, then, a population centric approach "focuses on the population, seeking to protect it from harm by — or interaction with — the insurgent, competing with the insurgent for influence and control at the grassroots level. Its basic assumption is that insurgency is a mass social phenomenon, that the enemy rides and manipulates a social wave consisting of genuine popular grievances, and that dealing with this broader social and political dynamic, while gaining time for targeted reforms to work by applying a series of tailored, full–spectrum security measures, is the most promising path to ultimately resolve the problem" (Kilcullen, 2009). In other words, a population centric approach to counterinsurgency requires convincing the civilian population that its interests are best served by the counterinsurgents, not the insurgents.

5.1.1.2 Enemy–centric

In Kilcullen's typology, the alternative to the population–centric approach to insurgency and counterinsurgency is the enemy–centric approach. Often maligned as the misapplication of conventional tactics to irregular war — see, for example, much of the work by Jonathan Nagl, though he doesn't use this terminology (Nagl, 2005) — the enemy–centric approach, like the population–centric approach, begins with the acknowledgement that irregular and conventional warfare are in many ways distinct. In fact, advocates of

the enemy–centric approach often even agree with their counterparts about what factors make them distinct. For example, C.E. Callwell's late–19th century, enemy–centric text *Small Wars* — among the first extended treatments of insurgency and counterinsurgency as essentially different from conventional warfare — focuses at great length on the unique advantages of insurgents, noting their speed and flexibility, and the advantages afforded the insurgents by the support of the native population, such as what he calls "treachery on the part of ostensibly neutral bodies or tribes." In other words, he acknowledges that what are now called "population centric" tactics are available to insurgents and are, in fact, their greatest advantage in warfare (Callwell, 1996).

But according to Callwell, these advantages enjoyed by insurgents require an approach to counterinsurgency nearly opposite that proposed by population–centric theorists. This is because counterinsurgents are, Callwell suggests, essentially incapable of fighting according to the same tactics as insurgents. Recall from the discussion above that Galula contends that regular warfare is typically the final phase of a successful insurgency — enacted, in part out of necessity, once the insurgents' support has reached critical mass. Galula seems even to imply that once the insurgent army is of a certain size it should no longer even be capable of employing the tactics of irregular warfare. Callwell essentially applies the same logic to the counterinsurgent (Callwell, 1996).

5.1.2 Antecedents

The debate over the proper practice of counterinsurgency contains many echoes of earlier debates surrounding the strategic use of air power in the infancy of aerial combat. Beginning with H.G. Wells's *The War in the Air*, early attempts to imagine how warfare might be changed by the introduction of aircraft focused on how it might be used to strategically manipulate civilian populations to turn on their political leaders. This would appear to have been the purpose of the Italian airmen who took to dropping hand grenades on noncombatants during the Italo-Turkish war of 1911–12, perhaps the first instance of what became known as "strategic bombing." This turned out to be a losing strategy at the time, "helping to drive people into the insurgents' arms [...] and the longer the war, the more the Italians themselves tended to replace [grenades] with leaflets that called upon the enemy to surrender" (van Creveld, 2011).

The Italian general Giulio Douhet is the most well–known advocate of "strategic bombing." Douhet wrote that "no longer can areas exist in which life can be lived in safety and tranquility, nor can that battlefield any longer be limited to actual combatants." Success in modern warfare, according to Douhet, thus required focusing attacks on "the most vital civilian centers [which would] spread terror through the nation and quickly break down [the enemy's] material and moral resistence" (van Creveld, 2011). Despite obvious differences, "strategic bombing" is in many ways analogous to "population–centric" approaches to counterinsurgency. In this case, "winning over" the

civilian population involves convincing them not to switch sides, but rather to force their own governments to make peace.

Douhet's counterparts in the US and the UK — Billy Mitchell and Hugh Trenchard, respectively — were somewhat more circumspect on the subject of aerially targeting civilians, and, interestingly, Germany's *Luftwaffe* rejected the tactic entirely (initially, at least). The Second World War should have been the ultimate test case for and against the "strategic bombing" advocated by Douhet, and the unconditional surrender of German and Japanese forces that followed soon after the devastation of many of their major cities by Allied air forces would appear to confirm its efficacy. Yet there is still some debate about whether the connection between "strategic bombing" and victory is as clear as the Allied forces' success would seem to imply. van Creveld (2011), for example, writes that in Germany

> bombing certainly did not improve morale of those on the receiving side. [...] Yet it did not bring the German people to the point where the ceased to resist, let alone made them rise against their government. Even in the face of infernos such as Hamburg, social cohesion proved much stronger than anybody had expected. Streets were quickly cleared of rubble, essential services restored. Moreover, bombing gave German fighter pilots something to fight and, if necessary, die for.

Here, then, was another opportunity for belligerents to "agree to disagree."

5.1.3 Practical implications of disagreement

If it is true that the victor in insurgencies and counterinsurgencies — and war in general — can be determined in some part by civilian populations, and that both the influence of civilian populations on a conflict's outcomes and the influence of belligerents on civilian populations are difficult to observe and a matter of significant dispute, this creates a number of practical difficulties for belligerents. To clarify these ideas, we can think of the role of civilians in terms of the complications it creates for understanding irregular and regular warfare within the framework of what are commonly called "rationalist" theories of war. I will focus on three of these complications – one, that adversaries may disagree about what will be the true determinants of victory, two, that they may therefore disagree about their respective chances of winning future battles, and three, that they might also therefore disagree about who has won a battle that has already been fought. I will then develop a modeling approach that addresses all three.

5.1.4 Rationalist theories of war

The canonical rationalist model of war is the "costly lottery." Although it has fallen into some disfavor — often with good reason, as noted below — it is still a useful starting place, as it clarifies many of the problems that must be dealt with in understanding how wars begin and end.

In the costly lottery representation of conflict (Fearon, 1995), both parties have respective probabilities p and $1 - p$ of winning the war, some

value W they assign to winning, and they must pay some costs k_1 and k_2 to fight. Fighting such a war would thus have an expected value of $p \times W - k_1$ and $(1-p) \times W - k_2$.[3] Once the war has been fought, the victor receives W minus its cost of fighting, and the loser receives zero minus its cost of fighting.

In the absence of additional complications, states under these circumstances should always prefer a negotiated settlement to absolute war, because there are any number of mutually agreeable divisions of W that they would prefer to fighting (Fearon, 1995). State 1 should prefer any share larger than $p \times W - k_1$ to war, and state 2 should be willing to give state 1 any share less than $p \times W - k_2$ rather than go to war. Put another way, $k_1 + k_2$ represent the overall value lost by fighting — the inefficiencies of war — and both states should prefer finding some way of keeping and dividing up this value rather than fighting and losing it. The larger the values of k_1 and k_2 are, the larger the range of settlements that both states prefer to war, and as long as k_1 and k_2 are greater than zero — which we can assume given the loss of lives and treasure that war necessarily entails — some range of mutually agreeable settlements will always exist. Given this range of mutually agreeable divisions of W, why would such a war ever occur?

[3]Note that I have simplified the usual version of a costly lottery somewhat, by conflating the share of the object of war each side receives with the subjective value that each side assigns to it. Under most circumstances, the logical ramifications of my simplified version will be the same as those of the more complicated canonical model.

5.1.5 Overestimating the probability of future success

One answer to this question that has inspired much of the contemporary literature on conflict is mutual optimism (Fearon, 1995). It may be that one or both sides overestimate their probabilities of winning, assigning subjective values to p — p_1 and $1-p_2$ — such that $p_1+1-p_2 > 1$ ($p_1-p_2 > 0$). The larger are p_1 and $(1-p_2)$, the smaller are the range of divisions of W that both would prefer to war. Once the difference between $p_1 \times W$ and $(1-p_2) \times W$ is greater than $k_1 + k_2$, the range of possible settlements preferred to war disappears entirely, such that $p_1 \times W - k_1$and $(1-p_2) \times W - k_2$ are greater than any division of W that the other side would agree to.

Note that if we understand mutual optimism as a result of adversaries having differential access to all of the information relevant to determining the likely victor, then such mutual optimism on its own is not enough to lead to war. If an asymmetric distribution of information were the only problem, states could avoid war entirely by directly communicating to each other whatever factors they believe give them an advantage, thus disabusing each other of their respective incorrect assumptions about future success. Typically, analyses that rely upon this conceptualization of mutual optimism within the context of a "costly lottery" model of war suggest that adversaries are unable to communicate in this manner due to incentives to misrepresent their own capabilities: If his share of the final settlement rises with his chances of winning p, then a belligerent should want to convince his adversary that p is very large — even if it isn't true. That they might lie makes any attempt by belligerents

to communicate private information inherently untrustworthy, thus making it difficult to overcome mutual optimism.

Smith and Stam have suggested an alternative approach. They reconstruct the rationalist model of conflict without Harsanyi's "common priors assumption," such that adversaries with access to the same information can hold different beliefs about the significance of this information. They can, in essence "agree to disagree." Some worry that such a departure from the canonical approach to rational choice theorizing is not only unnecessary but perhaps dangerous, leading to intractability and incoherence, ultimately undermining the ability to construct useful, causal explanations.[4] But Smith and Stam show that given careful consideration of what it is possible for belligerents to agree to disagree *about*, novel and valid explanations of wartime behavior can be constructed by jettisoning the common priors assumption. In the next sections, I will elaborate Smith and Stam's model, and then suggest another way that belligerents might agree to disagree.

5.1.6 The belligerent's ruin

In Smith and Stam's characterization, war is modeled as a variant of the "gambler's ruin" problem posed by Blaise Pascal and solved by Christiaan

[4]Fey and Ramsay make this argument quite adamantly, insisting on the necessity of the convention of treating differences in beliefs as, essentially, differences in available information (as in Fearon's and others' models of mutual optimism, discussed above). Strangely, in another article published at roughly the same time, Fey and Ramsay argue that mutual optimism based on informational asymmetries cannot ever lead to war, calling into question the usefulness of the very beliefs–as–information approach they advocate in their other article.

Huygens. In the "gambler's ruin" problem, two players with finite stakes make a series of bets that continue until one has lost and the other has won everything. The "problem" as defined by Pascal is how to determine the probability of each player winning the entire stake, and how long it will take them to reach one or the other of these two absorbing states. Huygens has shown that given a constant probability of winning, a constant amount won or lost with each round, and the players' initial bankrolls, it is possible to calculate each of these expected values. Given the total stake $X \in \mathbb{Z}^*$ (i.e., the sum of both players' bankrolls), player 1's probability of winning $p : 0 < p < 1$, and player 1's bankroll $x : x \in \mathbb{Z}^*$, and a bet of 1 unit for every round, 1's probability of losing everything is

$$\lambda(p,x) = \begin{cases} \frac{(\frac{1-p}{p})^x - (\frac{1-p}{p})^X}{1-(\frac{1-p}{p})^X} & \text{if } p \neq 1/2 \\ \\ \frac{(X-x)}{X} & \text{if } p = 1/2 \end{cases} \tag{5.1}$$

and player 2's probability of losing will be $1 - \lambda$. The expected number of periods until 1 or 2 wins everything is

$$E[R|p,x] = \begin{cases} \frac{1}{1-2p}[x - X(\frac{1-(\frac{1-p}{p})^x}{1-(\frac{1-p}{p})^X})]] & \text{if } p \neq 1/2 \\ \\ x(X-x) & \text{if } p = 1/2 \end{cases} \tag{5.2}$$

In Smith and Stam's adaptation, adversaries with finite "forts" (equivalent to the bankrolls in the gambler's ruin problem) fight a series of battles, each costing the belligerents K. However, unlike in the "gambler's ruin" problem, the purpose of fighting is the acquisition not only of forts, but also of some other prize that both sides value at V. Each can gain all of this prize by

depriving the other of all of her forts, and thus the ability to continue fighting. But at any time the adversaries can also choose to stop fighting by agreeing to some division of the prize. In each round t, before each battle, 1 offers 2 a division of the prize $b_t \in [0, V]$. If 2 accepts, then 1 gets $V - b_t - tK$ and 2 gets $b_t - tK$. If 2 rejects 1's offer, then they fight another battle. Each battle ends with the winner gaining and the loser losing one fort. Fighting can continue until one has lost and the other has won all of the forts. Given some known probability p of winning each battle and some known division of X forts, Smith and Stam use Huygen's result to determine the probability of each adversary reaching her preferred absorbing state (i.e., where the other has lost all of her forts), and how long it is likely to take to reach that state. Then, given the cost of fighting K and the value of the prize V, the expected values of fighting to the finish for belligerents 1 and 2 are

$$\begin{aligned} E_1 &= E[U_1(\text{fight indefinitely}|p, x)] \\ &= V \times (1 - \lambda(p, x)) - K \times E[R|p, x] \end{aligned}$$

and

$$\begin{aligned} E_2 &= E[U_2(\text{fight indefinitely}|p, x)] \\ &= V \times (\lambda(p, x)) - K \times E[R|p, x] \end{aligned}$$

respectively. If we maintain the common priors assumption, then both sides agree on the value of p, such that $E_1 + E_2 < V$, and there is a range of divisions of V (the size of which will be determined by K) between E_1 and E_2 that both

will prefer to war, much like in the canonical costly lottery model of mutual optimism.

But Smith and Stam instead assume that (given the same information) both sides can disagree about the probability of winning each battle. They represent adversaries' different beliefs about their probability of winning as random variables, beta distributed with parameters α_1, β_1 and α_2, β_2 for belligerents 1 and 2 respectively. Thus, given its beliefs about p, belligerent 1 expects to win a fight to the finish with probability

$$L_1(\alpha_1, \beta_1, x) = \int_0^1 (1 - \lambda(p, x)) f(p; \alpha_1, \beta_1) dp \tag{5.3}$$

and that the fight will last

$$D_1(\alpha_1, \beta_1, x) = \int_0^1 E[R|p, x] f(p; \alpha_1, \beta_1) dp \tag{5.4}$$

periods. Belligerent 2 similarly expects to win a fight to the finish with probability

$$L_2(\alpha_2, \beta_2, x) = \int_0^1 (\lambda(p, x)) f(p; \alpha_2, \beta_2) dp \tag{5.5}$$

and that the fight will last

$$D_2(\alpha_2, \beta_2, x) = \int_0^1 E[R|p, x] f(p; \alpha_2, \beta_2) dp \tag{5.6}$$

periods. Here $f(p; \alpha, \beta))$ is the density function for the beta distribution

$$f(p; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{(\alpha-1)} (1 - p)^{(\beta-1)} \tag{5.7}$$

where $\Gamma(.)$ is the gamma function, such that

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy \tag{5.8}$$

Thus a fight to the finish will be worth

$$\begin{aligned} E_1 &= E[U_1(\text{fight indefinitely}|\alpha_1, \beta_1, x)] \\ &= V \times (1 - L(\alpha_1, \beta_1, x)) - K \times D(\alpha_1, \beta_1, x) \end{aligned} \tag{5.9}$$

for 1 and thus 1 will not propose any division of V less than E_1. Belligerent 2 will likewise expect a fight to the finish to be worth

$$\begin{aligned} E_2 &= E[U_2(\text{fight indefinitely}|\alpha_2, \beta_2, x)] \\ &= V \times (L(\alpha_2, \beta_2, x)) - K \times D(\alpha_2, \beta_2, x) \end{aligned} \tag{5.10}$$

and will not accept any division of V less than E_2.

However, if they begin with mutually optimistic beliefs about p, 1 is willing to offer at most $M_1 = V - E_1$ and 2 likewise is willing to offer $M_2 = V - E_2$. When $E_2 > M_1$ or, similarly, when $E_1 > M_2$ (that is, when $V < E_1 + E_2$), each side's expected value of fighting to the finish is greater than the most generous offer the other is willing to make, so they will continue to fight.

Following Wittman's and Wagner's suggestion that war is a means for belligerents to learn about their likelihood of winning, Smith and Stam propose that 1 and 2 can update their respective beliefs α_1, β_1, α_2, and β_2 about p after each battle they fight via Bayes's rule. Thus, if at time t, belligerent 1 has won w_t and lost l_t battles, then $\alpha_{1t} = \alpha_{10} + w_t$ and $\beta_{1t} = \beta_{10} + l_t$. Because

$$w_t = \frac{t + X_t - X_0}{2} \tag{5.11}$$

and

$$l_t = \frac{t - X_t + X_0}{2} \tag{5.12}$$

both players' beliefs about p can be expressed in terms of time t and the state variable X_t, this is a Markovian random walk model of conflict. That is to say that the process it describes is essentially memoryless, such that beliefs at time $t+1$ depend only on beliefs at t, plus some stochastic element (determined by the true value of p, unknown to both belligerents). In other words, to describe belligerents' current beliefs about p does not require that we know the entire history of their beliefs, but only their most recently held beliefs and whatever event has caused them to revise them.

With this mechanism for updating beliefs in place, Smith and Stam show that 1's and 2's beliefs about p converge over time as they fight. They offer a formal proof, but the convergence result is evident because, as long as the belligerents avoid either absorbing state, the series p_1 and p_2 by construction both converge to .5 as t gets very large (and if they reach an absorbing state, then convergence is no longer necessary). Once p_1 and p_2 are in the same neighborhood — the boundaries of which are determined by K — then 1 and 2 will prefer a deal to another round of fighting. Given that K defines the size of the range of settlements 1 and 2 will prefer to war, as K gets smaller, there will be fewer settlements both belligerents prefer to war. Thus, belligerents' beliefs will converge more quickly when the cost of fighting is high than they will when the cost of fighting is low.

5.2 Divergent beliefs in insurgency and counterinsurgency

As already noted above, one feature of Smith and Stam's model is that beliefs always move closer together with each battle, and never move apart — this feature of the model, in fact, drives its convergence result. Substantively, this is because Smith and Stam assume that after the onset of war, belligerents have equal access to the same new information and interpret all of it the same way. Fighting thus provides both belligerents with immediate and identical evidence in support of one's beliefs or the other's: If one side prevails in battle, then its optimism about the determinants of victory is vindicated, and the other side's undermined. Similarly, if it loses, its optimism is undermined, and the other side's vindicated. Thus their beta distributed beliefs about their probabilities of winning, while initially different, update according to the same mechanism, ultimately converging.

But, as discussed above, the ultimate success of an insurgency or counterinsurgency may be determined to some degree by outcomes outside of the battlefield — e.g. the response of the civilian population to the battle and the extent to which this response contributes to future battlefield successes. Because they are not immediately observed by belligerents, these effects must be estimated, and different initial beliefs about the determinants of success may lead to different estimates. These estimates will then inform any revisions the belligerents make to their beliefs after battle. A belligerent who believes that counterinsurgency can be won primarily through enemy–centric means

might view a battle won through the overwhelming use of force as an unqualified success for the counterinsurgent, moving them closer to ultimate victory (Callwell, 1996). But another might focus on the collateral casualties from the same battle, and imagine that the long-term effects of the accompanying loss of support for the counterinsurgent from the civilian population will outweigh any gains made on the battlefield (Galula, 2006). Similarly, a belligerent who believes that counterinsurgency can be won primarily through population–centric means might view a battle fought with careful timing and precision, avoiding all civilian casualties, as an unqualified success for the counterinsurgent (Galula, 2006). But another might focus on the opportunities such caution provides the insurgent to regroup, recruit, regain the initiative, and continue the fight on its own terms, thus reducing the counterinsurgent's likelihood of success in future battles (Callwell, 1996). Thus, differences in initial beliefs may affect how these beliefs are revised in the face of new information, and these differences might in the short run cause belligerents' beliefs to diverge after a battle rather than converge.

To allow for disagreement over battlefield outcomes, first I treat the basis for such disagreement as being over the state of the world — i.e., whether the belligerents are in a world where enemy–centric counterinsurgency is better or a world where population–centric counterinsurgency is better. I define "better" here narrowly: While some battles will have unambiguous outcomes, such that regardless of the counterinsurgent's strategy both counterinsurgent and insurgent will agree on the outcome, the rest will have ambiguous outcomes

— due, for example, to the proximity of the battle to civilian populations and the likelihood of accompanying "collateral" casualties. Where a battle has an ambiguous outcome, if the belligerents are in the world where enemy–centric counterinsurgency is "better," this means that counterinsurgents will win when they have chosen to fight according to an enemy–centric strategy, but lose when they have chosen to fight according to a population–centric strategy. Likewise, if they are in the world where population–centric counterinsurgency is "better," this means they will win when they have chosen to fight according to a population–centric strategy, and lose when they have chosen to fight according to an enemy–centric strategy. I denote two state variables, representing proximity to an absorbing state in each of these two worlds. Next, I assign to each belligerent a set of beliefs about the probability of winning conditional on the state of the world, the counterinsurgent's choice of strategy, beliefs about the probability of being in one world or the other, and proximity to an absorbing state in each world. Finally, belligerents update their conditional probabilities based on which world they believe themselves to be in, such that if they disagree on the latter, they may also disagree on the former. I will explain each of these steps in greater detail.

As in Smith and Stam's model, the insurgent and counterinsurgent can gain all of stake V by depriving the other of all of her forts, and thus the ability to continue fighting. But at any time the adversaries can also choose to stop fighting by agreeing to some division of the stake. In each round t, before each battle, the counterinsurgent 1 offers the insurgent 2 a division of

the prize $b_t \in [0, V]$. If 2 accepts, then 1 gets $V - b_t - tK$ and 2 gets $b_t - tK$. If 2 rejects 1's offer, then they fight another battle. Thus, based on their beliefs, counterinsurgents make an offer and choose between two possible strategies: adopting a population-centric approach G (G is for Gallula) if the insurgents reject their offer or adopting an enemy-centric approach C (C is for Callwell) if the insurgents reject their offer. Likewise, given their own beliefs, insurgents choose between fighting and agreeing to the the counterinsurgent's settlement. Insurgents and counterinsurgents choose their strategies based on the expected utility of doing so, which I describe mathematically below. Note that this model can be understood as a special case of a more general model where insurgents also can decide between population centric (guerilla) and enemy centric (regular) warfare. Here I assume for the sake of greater simplicity that the insurgents' capacity is always sufficiently limited, as in what Galula identifies as the initial phases of all insurgencies, such that regular warfare is not a practical consideration.

Assuming either *might* be true, belligerents 1 and 2 begin with beliefs p_1^W and p_2^W respectively that they are in a world where the counterinsurgents' choice of G wins battles with ambiguous outcomes, and $1 - p_1^W$ and $1 - p_2^W$ that they are in a world where the counterinsurgents' choice of C wins battles with ambiguous outcomes. Each belligerent also begins with beliefs about the likelihood of winning in each scenario — the counterinsurgents' choice of G where G is best, of G where C is best, of C where C is best, and of C where G is best — four in total. Note that because I am interested in convergence

and divergence of beliefs, I will express beliefs for both belligerents in terms of their expectations about belligerent 1's (i.e. the counterinsurgents') likelihood of winning, such that belligerent 2's beliefs will be about their own likelihood of *losing*. The belligerents' first set of beliefs consists of their expected probabilities of the counterinsurgents' winning a battle if the counterinsurgents choose G and they are in a world where G is the best choice, p_1^{GG} and p_2^{GG}. Second is their expected probabilities of the counterinsurgent winning a battle if the counterinsurgents choose G and they are in a world where C is the best choice, p_1^{GC} and p_2^{GC}. Third is the expected probability of the counterinsurgents' winning a battle if the counterinsurgents choose C and they are in a world where C is the best choice, p_1^{CC} and p_2^{CC}. Fourth is the expected probability of the counterinsurgents' winning a battle if the counterinsurgents choose C and they are in a world where G is the best choice, p_1^{CG} and p_2^{CG}. To allow for Bayesian updating of beliefs, as in Smith and Stam's model, p_1^W and p_2^W, p_1^{GG} and p_2^{GG}, p_1^{GC} and p_2^{GC}, p_1^{CC} and p_2^{CC}, and p_1^{CG} and p_2^{CG} are all beta distributed, each with its own α and β shape parameters.[5]

As in Smith and Stam's model, each belligerent also begins with a finite number of "forts," and with each battle won (lost), she gains (loses) one fort.

[5] A note on the notation of probabilities and the parameters describing their distributions. The subscript always indicates the belligerent — 1 or 2. Where there are two superscripts, the first superscript indicates the strategy, and the second superscript indicates which world they may be in. Thus p_2^{GC} is the probability 2 assigns to winning given the counterinsurgents' choice of strategy G if they are in the world where C is the better strategy. Where there is one subscript, it indicates which world they may be in. Thus, p_2^W is the probability 2 assigns to their being in a world where the counterinsurgents' choice of G wins battles with ambiguous outcomes.

Once a belligerent loses or gains all of the forts, she stops fighting. But because belligerents can disagree about which world they are in, and the definition of success can be different in each world, I assign two state (or "fort") variables, X^C if they are in a world where the enemy–centric approach C is better, and X^G if they are in a world where the population–centric approach G is better. I assume that belligerents agree about how many battles they have each won and lost in each world, i.e. X^C and X^G are public information and the same for both belligerents.

Given these state variables X^C and X^G, the following expressions represent the belligerents' expectations about fighting.

5.2.1 Counterinsurgents' beliefs about choosing C

First I will describe expectations associated with the counterinsurgent's choice C. Here, given initial beliefs α_1^{CC} and β_1^{CC}, 1's expected probability of winning a fight to the finish in world C is

$$\begin{aligned} L_1^{CC}(.) &= L_1^{CC}(\alpha_1^{CC}, \beta_1^{CC}, X^C) \\ &= \int_0^1 (1 - \lambda(p, X^C)) f(p; \alpha_1^{CC}, \beta_1^{CC}) dp \end{aligned} \tag{5.13}$$

and 1 expects a fight to the finish to last

$$\begin{aligned} D_1^{CC}(.) &= D_1^{CC}(\alpha_1^{CC}, \beta_1^{CC}, X^C) \\ &= \int_0^1 E[R|p, X^C] f(p; \alpha_1^{CC}, \beta_1^{CC}) dp \end{aligned} \tag{5.14}$$

additional periods.

Given initial beliefs α_1^{CG} and β_1^{CG}, 1's expected probability of winning a fight to the finish in world G is

$$\begin{aligned} L_1^{CG}(.) &= L_1^{CG}(\alpha_1^{CG}, \beta_1^{CG}, X^G) \\ &= \int_0^1 (1 - \lambda(p, X^G)) f(p; \alpha_1^{CG}, \beta_1^{CG}) dp \end{aligned} \tag{5.15}$$

and 1 expects a fight to the finish to last

$$\begin{aligned} D_1^{CG}(.) &= D_1^{CG}(\alpha_1^{CG}, \beta_1^{CG}, X^G) \\ &= \int_0^1 E[R|p, X^G] f(p; \alpha_1^{CG}, \beta_1^{CG}) dp \end{aligned} \tag{5.16}$$

additional periods.

Given initial beliefs α_1^W and β_1^W that they are in world G, 1's expected probability of winning a fight to the finish is

$$\begin{aligned} L_1^C(.) &= L_1^C(\alpha_1^W, \beta_1^W, \alpha_1^{CC}, \beta_1^{CC}, \alpha_1^{CG}, \beta_1^{CG}, X_C, X_G) \\ &= \int_0^1 (p L_1^{CG}(.) + (1-p) L_1^{CC}(.)) f(p; \alpha_1^W, \beta_1^W) dp \end{aligned} \tag{5.17}$$

corresponding to Smith and Stam's equation 5.3 above, and for the fight to last

$$\begin{aligned} D_1^C(.) &= D_1^C(\alpha_1^W, \beta_1^W, \alpha_1^{CC}, \beta_1^{CC}, \alpha_1^{CG}, \beta_1^{CG}, X_C, X_G) \\ &= \int_0^1 (p D_1^{CG}(.) + (1-p) D_1^{CC}(.)) f(p; \alpha_1^W, \beta_1^W) dp \end{aligned} \tag{5.18}$$

periods, corresponding to Smith and Stam's equation 5.4 above.

Thus, the expected value of choice C for the counterinsurgents (1) is

$$\begin{aligned} E_1^C &= E[U_1(C|\alpha_1^W, \beta_1^W, \alpha_1^{CC}, \beta_1^{CC}, \alpha_1^{CG}, \beta_1^{CG}, X_C, X_G)] \\ &= V \times L_1^C(.) - K \times D_1^C(.) \end{aligned}$$

corresponding to Smith and Stam's equation 5.9 above.

5.2.2 Insurgents' beliefs about counterinsurgents choosing C

Given initial beliefs α_2^{CC} and β_2^{CC}, if the counterinsurgents choose C, the insurgents' (2's) expected probability of *losing* (i.e. of the counterinsurgents winning) a fight to the finish in world C is

$$\begin{aligned} L_2^{CC}(.) &= L_2^{CC}(\alpha_2^{CC}, \beta_2^{CC}, X^C) \\ &= \int_0^1 (1 - \lambda(p, X^C)) f(p; \alpha_2^{CC}, \beta_2^{CC}) dp \end{aligned} \tag{5.19}$$

and the insurgents (2) expect a fight to the finish to last

$$\begin{aligned} D_2^{CC}(.) &= D_2^{CC}(\alpha_2^{CC}, \beta_2^{CC}, X^C) \\ &= \int_0^1 E[R|p, X^C] f(p; \alpha_2^{CC}, \beta_2^{CC}) dp \end{aligned} \tag{5.20}$$

additional periods.

Given initial beliefs α_2^{CG} and β_2^{CG}, the insurgents' (2's) expected probability of *losing* (i.e. of the counterinsurgents winning) a fight to the finish in world G is

$$\begin{aligned} L_2^{CG}(.) &= L_2^{CG}(\alpha_2^{CG}, \beta_2^{CG}, X^G) \\ &= \int_0^1 (1 - \lambda(p, X^G)) f(p; \alpha_2^{CG}, \beta_2^{CG}) dp \end{aligned} \tag{5.21}$$

and the insurgents (2) expect a fight to the finish to last

$$\begin{aligned} D_2^{CG}(.) &= D_2^{CG}(\alpha_2^{CG}, \beta_2^{CG}, X^G) \\ &= \int_0^1 E[R|p, X^G] f(p; \alpha_2^{CG}, \beta_2^{CG}) dp \end{aligned} \tag{5.22}$$

additional periods.

Given initial beliefs α_2^W and β_2^W that they are in world G, the insurgents' (2's) expected probability of *losing* (i.e. of the counterinsurgents winning) a fight to the finish is

$$\begin{aligned} L_2^C(.) &= L_2^C(\alpha_2^W, \beta_2^W, \alpha_2^{CC}, \beta_2^{CC}, \alpha_2^{CG}, \beta_2^{CG}, X_C, X_G) \\ &= \int_0^1 (pL_2^{CG}(.) + (1-p)L_2^{CC}(.))f(p; \alpha_2^W, \beta_2^W)dp \end{aligned} \tag{5.23}$$

corresponding to Smith and Stam's equation 5.3 above, and they expect the fight to last

$$\begin{aligned} D_2^C(.) &= D_2^C(\alpha_2^W, \beta_2^W, \alpha_2^{CC}, \beta_2^{CC}, \alpha_2^{CG}, \beta_2^{CG}, X_C, X_G) \\ &= \int_0^1 (pD_2^{CG}(.) + (1-p)D_2^{CC}(.))f(p; \alpha_2^W, \beta_2^W)dp \end{aligned} \tag{5.24}$$

periods, corresponding to Smith and Stam's equation 5.4 above.

Thus, the expected value of fighting for the insurgent (2) given the counterinsurgents' (1's) choice of C is

$$\begin{aligned} E_2^C &= E[U_2(C|\alpha_2^W, \beta_2^W, \alpha_2^{CC}, \beta_2^{CC}, \alpha_2^{CG}, \beta_2^{CG}, X_C, X_G)] \\ &= V \times (1 - L_2^C(.)) - K \times D_2^C(.) \end{aligned}$$

corresponding to Smith and Stam's equation 5.9 above.

5.2.3 Counterinsurgents beliefs about choosing G

Next, I will describe expectations associated with the counterinsurgent's choice G. Here, given initial beliefs α_1^{GG} and β_1^{GG}, 1's expected proba-

bility of winning a fight to the finish in world G is

$$\begin{aligned} L_1^{GG}(.) &= L_1^{GG}(\alpha_1^{GG}, \beta_1^{GG}, X^G) \\ &= \int_0^1 (1 - \lambda(p, X^G)) f(p; \alpha_1^{GG}, \beta_1^{GG}) dp \end{aligned} \tag{5.25}$$

and 1 expects a fight to the finish to last

$$\begin{aligned} D_1^{GG}(.) &= D_1^{GG}(\alpha_1^{GG}, \beta_1^{GG}, X^G) \\ &= \int_0^1 E[R|p, X^G] f(p; \alpha_1^{GG}, \beta_1^{GG}) dp \end{aligned} \tag{5.26}$$

additional periods.

Given initial beliefs α_1^{GC} and β_1^{GC}, 1's expected probability of winning a fight to the finish in world C is

$$\begin{aligned} L_1^{GC}(.) &= L_1^{GC}(\alpha_1^{GC}, \beta_1^{GC}, X^C) \\ &= \int_0^1 (1 - \lambda(p, X^C)) f(p; \alpha_1^{GC}, \beta_1^{GC}) dp \end{aligned} \tag{5.27}$$

and 1 expects a fight to the finish to last

$$\begin{aligned} D_1^{GC}(.) &= D_1^{GC}(\alpha_1^{GC}, \beta_1^{GC}, X^C) \\ &= \int_0^1 E[R|p, X^C] f(p; \alpha_1^{GC}, \beta_1^{GC}) dp \end{aligned} \tag{5.28}$$

additional periods.

Given initial beliefs α_1^W and β_1^W that they are in world G, 1's expected probability of winning a fight to the finish is

$$\begin{aligned} L_1^G(.) &= L_1^G(\alpha_1^W, \beta_1^W, \alpha_1^{GG}, \beta_1^{GG}, \alpha_1^{GC}, \beta_1^{GC}, X_C, X_G) \\ &= \int_0^1 (p L_1^{GG}(.) + (1-p) L_1^{GC}(.)) f(p; \alpha_1^W, \beta_1^W) dp \end{aligned} \tag{5.29}$$

corresponding to Smith and Stam's equation 5.3 above, and for the fight to last

$$\begin{aligned} D_1^G(.) &= D_1^G(\alpha_1^W, \beta_1^W, \alpha_1^{GG}, \beta_1^{GG}, \alpha_1^{GC}, \beta_1^{GC}, X_C, X_G) \\ &= \int_0^1 (pD_1^{GG}(.) + (1-p)D_1^{GC}(.))f(p; \alpha_1^W, \beta_1^W)dp \end{aligned} \tag{5.30}$$

periods, corresponding to Smith and Stam's equation 5.4 above.

Thus, the expected value of choice G for the counterinsurgents (1) is

$$\begin{aligned} E_1^G &= E[U_1(C|\alpha_1^W, \beta_1^W, \alpha_1^{GG}, \beta_1^{GG}, \alpha_1^{GC}, \beta_1^{GC}, X_C, X_G)] \\ &= V \times L_1^C(.) - K \times D_1^C(.) \end{aligned}$$

corresponding to Smith and Stam's equation 5.9 above.

5.2.4 Insurgents' beliefs about counterinsurgents choosing G

Given initial beliefs α_2^{GG} and β_2^{GG}, 2's expected probability of losing a fight to the finish in world G is

$$\begin{aligned} L_2^{GG}(.) &= L_2^{GG}(\alpha_2^{GG}, \beta_2^{GG}, X^G) \\ &= \int_0^1 (1 - \lambda(p, X^G))f(p; \alpha_2^{GG}, \beta_2^{GG})dp \end{aligned} \tag{5.31}$$

and 2 expects a fight to the finish to last

$$\begin{aligned} D_2^{GG}(.) &= D_2^{GG}(\alpha_2^{GG}, \beta_2^{GG}, X^G) \\ &= \int_0^1 E[R|p, X^G]f(p; \alpha_2^{GG}, \beta_2^{GG})dp \end{aligned} \tag{5.32}$$

additional periods.

Given initial beliefs α_2^{GC} and β_2^{GC}, 2's expected probability of winning a fight to the finish in world C is

$$\begin{aligned} L_2^{GC}(.) &= L_2^{GC}(\alpha_2^{GC}, \beta_2^{GC}, X^C) \\ &= \int_0^1 (1 - \lambda(p, X^C)) f(p; \alpha_2^{GC}, \beta_2^{GC}) dp \end{aligned} \tag{5.33}$$

and 2 expects a fight to the finish to last

$$\begin{aligned} D_2^{GC}(.) &= D_2^{GC}(\alpha_2^{GC}, \beta_2^{GC}, X^C) \\ &= \int_0^1 E[R|p, X^C] f(p; \alpha_2^{GC}, \beta_2^{GC}) dp \end{aligned} \tag{5.34}$$

additional periods.

Given initial beliefs α_2^W and β_2^W that they are in world G, 2's expected probability of losing a fight to the finish is

$$\begin{aligned} L_2^G(.) &= L_2^G(\alpha_2^W, \beta_2^W, \alpha_2^{GG}, \beta_2^{GG}, \alpha_2^{GC}, \beta_2^{GC}, X_C, X_G) \\ &= \int_0^1 (p L_2^{GG}(.) + (1-p) L_2^{GC}(.)) f(p; \alpha_2^W, \beta_2^W) dp \end{aligned} \tag{5.35}$$

corresponding to Smith and Stam's equation 5.3 above, and for the fight to last

$$\begin{aligned} D_2^G(.) &= D_2^G(\alpha_2^W, \beta_2^W, \alpha_2^{GG}, \beta_2^{GG}, \alpha_2^{GC}, \beta_2^{GC}, X_C, X_G) \\ &= \int_0^1 (p D_2^{GG}(.) + (1-p) D_2^{GC}(.)) f(p; \alpha_2^W, \beta_2^W) dp \end{aligned} \tag{5.36}$$

periods, corresponding to Smith and Stam's equation 5.4 above.

Thus, the expected value for the insurgent (2) of the counterinsurgents' (1's) choice of G is

$$\begin{aligned} E_2^G &= E[U_2(C|\alpha_2^W, \beta_2^W, \alpha_2^{GG}, \beta_2^{GG}, \alpha_2^{GC}, \beta_2^{GC}, X_C, X_G)] \\ &= V \times (1 - L_2^C(.)) - K \times D_2^C(.) \end{aligned}$$

corresponding to Smith and Stam's equation 5.9 above.

5.2.5 How settlements are reached

The counterinsurgents (belligerent 1) are willing to offer at most

$$M_1 = V - max\{E_1^G, E_1^C\} \tag{5.37}$$

and 2 likewise is willing to offer

$$M_2 = V - max\{E_2^G, E_2^C\}. \tag{5.38}$$

When

$$max\{E_2^G, E_2^C\} > M_1 \tag{5.39}$$

or, similarly, when

$$max\{E_1^G, E_1^C\} > M_2 \tag{5.40}$$

that is, when

$$V < max\{E_1^G, E_1^C\} + max\{E_2^G, E_2^C\} \tag{5.41}$$

each side's expected value of fighting to the finish is greater than the most generous offer the other is willing to make, so they will continue to fight.

The mechanism by which the belligerents update their beliefs about the probability of winning is similar to Smith and Stam's, albeit complicated by the belligerents' ability to disagree about the outcome of a given battle. As in Smith and Stam's model, belligerents will update their beliefs about their probabilities of winning p_1^{GG}, p_2^{GG}, p_1^{GC}, p_2^{GC}, p_1^{CC}, p_2^{CC}, p_1^{CG}, and p_2^{CG} based on battlefield outcomes.

But where battle outcomes are ambiguous, belligerents can disagree about what those outcomes are. In these cases, a battle fought according to the enemy–centric approach, and scored as a win from an enemy–centric point of view, would be scored as a loss from the population–centric point of view. Similarly, a battle fought according to the population–centric approach, and scored as a win from a population–centric point of view, would be scored as a loss from the enemy–centric point of view. Thus, an ambiguous battle fought according to the enemy–centric strategy will be a win for the counterinsurgents according to the enemy–centric view but a loss according to the population–centric view. This would lead to an increase in α_1^{CC}, α_2^{CC}, α_1^{GC}, and α_2^{GC} (the parameters associated with an increase in the expected probability of winning if 1 and 2 are in the world where C is the best strategy) by one unit, and an increase in β_1^{GG}, β_2^{GG}, β_1^{CG}, and β_2^{CG} (the parameters associated with a decrease in the expected probability of winning in a world where G is the best strategy) by one unit. Likewise the "fort" tallies would also change: X^C would increase by one, and X^G would decrease by one. Similarly, an ambiguous battle fought according to the population–centric strategy will be

a win for the counterinsurgents according to the population–centric view but a loss according to the enemy–centric view. This would lead to an increase in α_1^{GG}, α_2^{GG}, α_1^{CG}, and α_2^{CG} (the parameters associated with an increase in the expected probability of winning if they are in the world where C is the best strategy) by one unit, and an increase in β_1^{CC}, β_2^{CC}, β_1^{GC}, and β_2^{GC} (the parameters associated with a decrease in the expected probability of winning in a world where G is the best strategy) by one unit. Likewise the "fort" tallies would also change: X^G would increase by one, and X^C would decrease by one.[6]

There is a separate mechanism for updating the α and β parameters associated with p_1^W and p_2^W — 1's and 2's respective beliefs that they are in a world where G is the better strategy. Because the perception of battlefield outcomes is essentially biased by prior beliefs about whether an enemy-centric or population-centric approach is better, any measure of battlefield success will be a poor indicator of whether these prior beliefs are correct. Updating the α and β parameters associated with p_1^W and p_2^W requires that belligerents

[6]Here I assume (or, more precisely, I assume that the belligerents assume) that if the counterinsurgent's pursuing an enemy–centric strategy gives an advantage to the insurgent, this advantage to the insurgent will persist even if the counterinsurgents realize their mistake and switch to a population-centric strategy. That is, they will have disadvantaged themselves in future battles regardless of whatever additional advantage is gained or lost by their future choices of strategy. Likewise, if the counterinsurgents' pursuing a population strategy disadvantages the insurgents, this disadvantage will persist even if the counterinsurgents don't realize this, and make the mistake of switching to an enemy-centric strategy. Again, they will have given themselves an advantage in future battles regardless of their future strategy. Thus, when a battle is fought according to the enemy–centric approach, and won according to the enemy–centric point of view, but lost according to the population–centric point of view, α_1^{GC} and α_2^{GC} will increase by one, while β_1^{CG} and β_2^{CG} will increase by one.

pay attention to some measure of long run success that both the enemy– and population–centric approaches would agree is valid, and about which there is little room for subjective disagreement between belligerents over outcomes.

One such measure is the frequency of insurgent attacks. Theorists of insurgency, population–centric counterinsurgency, and enemy–centric counterinsurgency all agree that the best way for insurgents to fight is via small, frequent skirmishes, rather than large scale, direct confrontation. The logic behind such an emphasis on small, frequent attacks is straightforward: Generally speaking, insurgents will not mobilize additional capacity in the manner that a regular army would, i.e. by increasing battalion strength in order to win more often and more decisively, as they will not be able to compete with the regular army of the counterinsurgents in direct confrontation. Rather, as their goal is not to defeat but rather to consistently undermine the legitimacy and authority of the larger military force, they are more likely to apply any newly acquired capacity to increasing the scope and frequency of attacks. In Mao's words, "the enemy is much stronger than we are, and it is true that we can hinder, distract, disperse, and destroy him only if we disperse our own forces." If this is the case, any increase in insurgent force capacity should lead to an increase in the frequency of insurgent initiated attacks, and likewise a decrease in capacity should lead to a decrease in the frequency of insurgent initiated attacks.

If it is true, as the population–centric approach maintains, that an enemy–centric approach creates recruitment opportunities for the insurgents

that exceed whatever attenuation of insurgent forces is achieved on the battlefield, then this increase in insurgent capacity should be reflected by an increase in the frequency of insurgent initiated attacks. Likewise, if it is true, as is implied by defenders of the enemy–centric approach, that the population–centric approach's overcommitment of troops to the provision of security for civilians creates too many opportunities for the insurgents to regroup, recruit, seize the initiative, and fight battles on their own terms, then this too should be reflected in an increase in the frequency of insurgent initiated attacks. In either case, an increase in the frequency of insurgent initiated attacks is evidence that the current approach is not working and that the alternative might be superior, whereas a decrease in the frequency of insurgent initiated attacks is evidence that the current approach is at the very least working and may even be the superior of the two approaches.

Thus, if counterinsurgents choose G and insurgent initiated attacks are seen to decrease (increase) over time, then the α (β) parameters associated with p_1^W and p_2^W will increase for *both* players. If counterinsurgents choose C and insurgent initiated attacks decrease (increase), then the β (α) associated with p_1^W and p_2^W will increase for *both* players.

The running tally of wins and losses are expressed in terms of time varying state variables X_t^C and X_t^G, as described above. So that all of the belligerents' expectations can be written as functions of Markovian random walk processes, the α and β parameters associated with p_1^W and p_2^W will update according to a third state variable, X_t^W, which increases when insurgent

attacks go up and decreases when they go down. Thus, at time t, the expected value of choice C for the counterinsurgents (belligerent 1) is

$$\begin{aligned} E_{1t}^{C} &= E[U_{1t}(C|X_t^C, X_t^G, X_t^W, t)] \\ &= V \times L_1^C \left(\begin{array}{c} \alpha_1^W + \frac{t+X_t^W-X_0^W}{2}, \beta_1^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_1^{CC} + \frac{t+X_t^C-X_0^C}{2}, \beta_1^{CC} + \frac{t-X_t^C+X_0^C}{2}, \\ \alpha_1^{CG} + \frac{t+X_t^G-X_0^G}{2}, \beta_1^{CG} + \frac{t-X_t^G+X_0^G}{2}, \\ X_t^C, X_t^G \end{array} \right) \\ &\quad -K \times D_1^C \left(\begin{array}{c} \alpha_1^W + \frac{t+X_t^W-X_0^W}{2}, \beta_1^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_1^{CC} + \frac{t+X_t^C-X_0^C}{2}, \beta_1^{CC} + \frac{t-X_t^C+X_0^C}{2}, \\ \alpha_1^{CG} + \frac{t+X_t^G-X_0^G}{2}, \beta_1^{CG} + \frac{t-X_t^G+X_0^G}{2}, \\ X_t^C, X_t^G \end{array} \right) \end{aligned} \tag{5.42}$$

the expected value of continuing to fight for the insurgent (2) when the counterinsurgents choose C is

$$\begin{aligned} E_{2t}^{C} &= E[U_2(C|X_t^C, X_t^G, X_t^W, t)] \\ &\quad V \times \left(1 - L_2^C \left(\begin{array}{c} \alpha_2^W + \frac{t+X_t^W-X_0^W}{2}, \beta_2^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_2^{CC} + \frac{t+X_t^C-X_0^C}{2}, \beta_2^{CC} + \frac{t-X_t^C+X_0^C}{2}, \\ \alpha_2^{CG} + \frac{t+X_t^G-X_0^G}{2}, \beta_2^{CG} + \frac{t-X_t^G+X_0^G}{2}, \\ X_t^C, X_t^G \end{array} \right) \right) \\ &\quad -K \times D_2^C \left(\begin{array}{c} \alpha_2^W + \frac{t+X_t^W-X_0^W}{2}, \beta_2^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_2^{CC} + \frac{t+X_t^C-X_0^C}{2}, \beta_2^{CC} + \frac{t-X_t^C+X_0^C}{2}, \\ \alpha_2^{CG} + \frac{t+X_t^G-X_0^G}{2}, \beta_2^{CG} + \frac{t-X_t^G+X_0^G}{2}, \\ X_t^C, X_t^G \end{array} \right) \end{aligned} \tag{5.43}$$

the expected value of choice G for the counterinsurgents (belligerent 1) is

$$\begin{aligned} E_{1t}^{G} &= E[U_1(G|X_t^C, X_t^G, X_t^W, t)] \\ & V \times L_1^G \left(\begin{array}{c} \alpha_1^W + \frac{t+X_t^W-X_0^W}{2}, \beta_1^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_1^{GG} + \frac{t+X_t^G-X_0^G}{2}, \beta_1^{GG} + \frac{t-X_t^G+X_0^G}{2}, \\ \alpha_1^{GC} + \frac{t+X_t^C-X_0^C}{2}, \beta_1^{GC} + \frac{t-X_t^C+X_0^C}{2}, \\ X_t^C, X_t^G \end{array} \right) \\ & -K \times D_1^G \left(\begin{array}{c} \alpha_1^W + \frac{t+X_t^W-X_0^W}{2}, \beta_1^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_1^{GG} + \frac{t+X_t^G-X_0^G}{2}, \beta_1^{GG} + \frac{t-X_t^G+X_0^G}{2}, \\ \alpha_1^{GC} + \frac{t+X_t^C-X_0^C}{2}, \beta_1^{GC} + \frac{t-X_t^C+X_0^C}{2}, \\ X_t^C, X_t^G \end{array} \right) \end{aligned} \tag{5.44}$$

and the expected value of continuing to fight for the insurgents (belligerent 2) when the counterinsurgents choose G is

$$\begin{aligned} E_{2t}^{G} &= E[U_2(G|X_t^C, X_t^G, X_t^W, t)] \\ & V \times \left(1 - L_2^C \left(\begin{array}{c} \alpha_2^W + \frac{t+X_t^W-X_0^W}{2}, \beta_2^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_2^{GG} + \frac{t+X_t^G-X_0^G}{2}, \beta_2^{GG} + \frac{t-X_t^G+X_0^G}{2}, \\ \alpha_2^{GC} + \frac{t+X_t^C-X_0^C}{2}, \beta_2^{GC} + \frac{t-X_t^C+X_0^C}{2}, \\ X_t^C, X_t^G \end{array} \right) \right) \\ & -K \times D_2^G \left(\begin{array}{c} \alpha_2^W + \frac{t+X_t^W-X_0^W}{2}, \beta_2^W + \frac{t-X_t^W+X_0^W}{2}, \\ \alpha_2^{GG} + \frac{t+X_t^G-X_0^G}{2}, \beta_2^{GG} + \frac{t-X_t^G+X_0^G}{2}, \\ \alpha_2^{GC} + \frac{t+X_t^C-X_0^C}{2}, \beta_2^{GC} + \frac{t-X_t^C+X_0^C}{2}, \\ X_t^C, X_t^G \end{array} \right). \end{aligned} \tag{5.45}$$

Unlike X^G and X^C, I assume no absorbing states associated with X^W. Because of this, belligerent's beliefs p_1^W and p_2^W do not by construction converge to .5, but rather, given sufficiently large t, converge to $\frac{k}{t}$, where t is the number of periods the belligerents have been fighting and k is the number of those periods that insurgent–initiated attacks have been observed to increase

when the counterinsurgent chooses C and decrease when the counterinsurgent chooses G. However, because p_1^{GG}, p_2^{GG}, p_1^{GC}, p_2^{GC}, p_1^{CC}, p_2^{CC}, p_1^{CG}, and p_2^{CG} all do converge to .5 as t gets very large, and $p^W \times .5 + (1 - p^W) \times .5 = .5$ for all $0 <= p^W <= 1$, then E_{1t}^C, E_{2t}^C, E_{1t}^G, and E_{2t}^G all converge to .5.

Let

$$E_{2t} = \begin{cases} E_{2t}^C & \text{if } E_{1t}^C > E_{1t}^G \\ E_{2t}^G & \text{if } E_{1t}^C < E_{1t}^G \\ \frac{E_{2t}^C + E_{2t}^G}{2} & \text{if } E_{1t}^C = E_{1t}^G \end{cases} \tag{5.46}$$

it follows, then, that for a sufficiently large t,

$$V \;\; \geq \;\; max\{0, E_{2t}\} + max\{0, E_{1t}^G, E_{1t}^C\} \tag{5.47}$$

where the $max\{\}$ operator returns the largest element in a given set of numbers.

That the insurgents and counterinsurgents stop fighting at this point, after 1 offers and 2 accepts $b = max\{0, E_{2t}\}$, can be established as follows. Once t is sufficiently large, the insurgent will accept any offer $b_{\geq} \geq max\{0, E_{2t}\}$ and reject any offer $b_{<} < max\{0, E_{2t}\}$. The cost of fighting makes rejection worth less than $b_{\geq}$ but is not high enough to make it worth less than $b_{<}$. Conversely, the counterinsurgents will be willing to make any offer $b_{\leq} \leq max\{0, E_{2t}\}$ rather than fight, but will never make an offer $b_{>} > max\{0, E_{2t}\}$. When $V = max\{0, E_{2t}\} + max\{0, E_{1t}^G, E_{1t}^C\}$, fighting will get the counterinsurgents $max\{0, E_{1t}^G, E_{1t}^C\} = V - max\{0, E_{2t}\}$, so they will not be willing to make an offer greater than $max\{0, E_{2t}\}$ (thus keeping for themselves less than $V - max\{0, E_{2t}\}$). But knowing that the insurgents will reject any offer less

than $max\{0, E_{2t}\}$, the counterinsurgents will offer $b = max\{0, E_{2t}\}$, which the insurgents will accept, thus ending fighting.

Smith and Stam have shown that, given that all of the possible random walks implied by their system of equations end in convergence, 1 can, via backwards induction, identify those nodes where she will prefer to make an offer that she knows 2 will reject (thus continuing fighting). 1's arrival at each node is essentially determined by a non-strategic Nature that determines (based on the true value of p, unknown to the belligerents) at each t whether a given X_t will equal $X_{t-1} + 1$ or $X_{t-1} - 1$ and whether 2 was willing to accept or reject 1's last offer. Thus, despite that 1's choice is not exactly discrete, the game can be modeled in a manner roughly equivalent to a normal form game tree, with branches representing moves by nature and 2's two possible responses to the best offers made by 1 — an admittedly enormous tree, but one that is nonetheless both finite and, given sufficient computing power and time, solvable. The result, then, is the ability to state for both players equilibrium strategies for every possible sequence of moves by nature.

My model is similarly constructed, although the implied game tree is even larger and more complicated. While the insurgent's (2's) choices are the same, the insurgent (1) must now not only make an offer but choose either C or G, and nature now has eight possible moves at each node: $\{X_t^W = X_{t-1}^W + 1, X_t^C = X_{t-1}^C + 1, X_t^G = X_{t-1}^G + 1\}$, $\{X_t^W = X_{t-1}^W + 1, X_t^C = X_{t-1}^C + 1, X_t^G = X_{t-1}^G - 1\}$, $\{X_t^W = X_{t-1}^W + 1, X_t^C = X_{t-1}^C - 1, X_t^G = X_{t-1}^G + 1\}$, $\{X_t^W = X_{t-1}^W + 1, X_t^C = X_{t-1}^C - 1, X_t^G = X_{t-1}^G - 1\}$, $\{X_t^W = X_{t-1}^W - 1, X_t^C =$

$X^C_{t-1}+1, X^G_t = X^G_{t-1}+1\}$, $\{X^W_t = X^W_{t-1}-1, X^C_t = Xt-1^C+1, X^G_t = X^G_{t-1}-1\}$, $\{X^W_t = X^W_{t-1}-1, X^C_t = X^C_{t-1}-1, X^G_t = X^G_{t-1}+1\}$, or $\{X^W_t = X^W_{t-1}-1, X^C_t = X^C_{t-1}-1, X^G_t = X^G_{t-1}-1\}$. As in Smith and Stam's model, it is possible to identify via backwards induction those nodes where 1 prefers to make an offer she knows 2 will reject, and thus to state for each player equilibrium strategies for every possible sequence of moves by nature.

In fact, most of the "branches" of this tree will look more or less exactly like what Smith and Stam describe. As long as ambiguous battles are relatively infrequent, belligerents' beliefs converge almost as quickly here as they do in Smith and Stam's model. But I am interested in one branch in particular — the one where *all* or *most* of the battles are ambiguous. Such ambiguity may be among the defining features of protracted conflicts, such as the wars currently being fought by the US in Iraq and Afghanistan, so it is important to contemplate what would be the strategic implications of sustaining such ambiguity.[7]

Even when all battles are ambiguous, in the long run beliefs and expectations converge. But in the short run other dynamics are possible, including

[7]A corollary is the question of *why* such ambiguity might be sustained — and, for that matter, *who* is sustaining it. One implication of this model may be that these ambiguities can be exploited to the advantage of the insurgent. If the counterinsurgents can be compelled repeatedly to act against their own best interests, by choosing to fight in a way and in a setting that actually undermines their future probability of success, then this may increase the final settlement offer they are willing to make. In this case it might be in the interests of the insurgents to *engineer* the kinds of ambiguity that would lead the counterinsurgents to make such mistakes — by, for example, choosing to fight in heavily populated areas where collateral casualties are more likely.

divergence. The model I've described is sensitive to the initial values of the beta functions' shape parameters (i.e. the αs and βs), making general statements about the conditions for short run divergence difficult. However, it is possible to compare the resulting dynamics of different models with plausible initial values. The remainder of this chapter will examine and compare a number of such plausible examples, and suggest some of the broader implications of these comparisons.

5.3 A model of bargaining, fighting and learning during the US led counterinsurgency in Iraq

There is a narrative common among many of the more prominent analyses of the Iraq war — both those that are critical of and sympathetic towards the enterprise itself — wherein occupying forces in Iraq compounded early blunders in their efforts to secure the peace by failing to quickly adapt to the changing realities of the conflict. According to this narrative, in the intervening years between invasion and surge, US forces in Iraq were implementing losing tactics and failing to adapt due, in large part, to an institutional prejudice against change (Diamond, 2005; Ricks, 2006; Hashim, 2006; Nagl, 2007; Robinson, 2008; Metz, 2008). Johnathan Nagl's is perhaps the most explicit statement of this conclusion, blaming the military's "unwillingness to internalize and build upon the lessons of Vietnam" for both its initial lack of preparation for combating insurgency and an institutional inability to "learn and adapt" to the changing nature of the conflict (Nagl, 2005, 2007).

While not necessarily incorrect, this is an incomplete picture of Iraq before the surge. Two pieces are missing from this account. First, that the adoption as doctrine of US Army Field Manual 3–24 (a.k.a. the Counterinsurgency Field Manual), the appointment of David Petraeus to Commanding General, Multi–National Force — Iraq, and the 2007 surge in US troop strength, were all *made possible* by earlier failures. While institutional inertia played a role in US forces' initial inability to contain growing violence and instability in occupied Iraq, the "surge" and the changes in tactics that accompanied it are evidence that US forces were ultimately able to learn, albeit slowly, from their mistakes and adapt to the realities of the conflict.

Second, what is meant by "learning" and "inertia" is underspecified in these critiques. About what, exactly, did counterinsurgent forces need to learn more quickly? What kinds of inertia, exactly, prevented them from doing so?

In the model presented above, there are two kinds of learning that are possible (and thus two corresponding kinds of inertia that might hinder learning). The first kind of learning is about the conditional probabilities of winning, p_1^{GG}, p_2^{GG}, p_1^{GC}, p_2^{GC}, p_1^{CC}, p_2^{CC}, p_1^{CG}, and p_2^{CG}. Substantively, these can be understood as beliefs about material capabilities — a function not only of weapons and troop strength, but also of the advantages and disadvantages afforded by terrain and the cooperation (or lack thereof) the local civilian population. The second kind of learning is about the state of the world, p_1^W and p_2^W. Substantively, these can be understood as beliefs about whether choice of strategy and/or tactics matters and, if so, which choice is better.

The beta distribution is well suited for modeling varying degrees of inertia. The parameters of a beta distribution determine the overall shape of its curve, such that the values of α and β for a given distribution of beliefs about p determine not only the expected value or mean of p, $E[p]$, but also its variance, $V[p]$, which can be understood substantively as a measure of certainty. Regardless of whether $\alpha = 2$ and $\beta = 4$ or $\alpha = 20$ and $\beta = 40$, $E[p]$ will be the same: $\frac{1}{3}$. But $V[p]$ will be equal to .03 in the first case, and .004 in the second, a significant difference. Substantively, certainty should be related to sensitivity to new information — the more certain one is, the less of an influence new information will have on one's beliefs, an intuition that the beta function appropriately represents. For example, a loss for a belligerent in the model above will entail increasing the β of the corresponding p by 1. In the first case, where $\alpha = 2$ and $\beta = 4$, this will mean $E[p]$ moves from 1/3 (.333) to 2/7 (.290), a not insignificant reduction in their expected likelihood of winning. In the second case, $\alpha = 20$ and $\beta = 40$, it will mean $E[p]$ moves from 1/3 (.333) to 20/41 (.328), a much smaller reduction. Thus, the larger are the shape parameters α and β of a belligerent's beliefs about p, the less sensitive these beliefs are to new information.

In other words, the relative orders of magnitude of the shape parameters α and β represent the relative inertia of the belligerents' corresponding beliefs. Where this inertia is relatively high (i.e. α and β are high), learning and adaptation are more difficult and slower. Where inertia is relatively low (i.e. α and β are low), learning and adaptation are easier and faster.

Given the two types of learning the model allows for, this suggests a number of possible scenarios. I focus on the following comparisons: First, inertia hindering learning about means and inertia hindering learning about strategy could both be roughly the same order of magnitude — i.e. both high or both low. Second, inertia hindering learning about means could be high while inertia hindering learning about strategy could be low. Finally, inertia hindering learning about means could be low while inertia hindering learning about strategy could be high. In each case, I assume an initial state of unconditional mutual optimism. That is, both the insurgent and counterinsurgent believe they have a high probability of winning regardless of strategy/world. They also disagree about which is the best strategy — the insurgent believes she is in a world where population–centric counterinsurgency is better and the counterinsurgent believes she is in a world where enemy–centric counterinsurgency is better. I assume furthermore that the insurgent is flexible (i.e. uncertain, or has low inertia) about both means and strategy — that is, the insurgent has starting values for each α and β parameter that are of a low order of magnitude. I assume that the insurgent, furthermore, is *right* about the state of the world. That is, after each battle, both α_1^W and α_2^W will go up. In all cases, X^C and X^G (the number of forts held by the counterinsurgent in worlds C and G respectively) both have initial values of 25, and N^C and N^G (the total number of forts) are both 50.

I discuss each case in turn. Tables 1 through 3 show how the expected probabilities of winning for both belligerents change during a series of ambigu-

ous battles in each scenario. Columns 1 through 4 in table 1 show how belligerents' beliefs about the counterinsurgent's probability of winning given strategy C or G change with each battle when the counterinsurgent is equally certain about means and strategy.[8] Columns 5 is the difference between the counterinsurgent's and the insurgent's beliefs about the counterinsurgent's probability of winning given C, and column 6 is the difference between the counterinsurgent's and the insurgent's beliefs about the counterinsurgent's probability of winning given G. Thus, for five rounds of fighting, the insurgent's and counterinsurgent's beliefs diverge, after which point they start a very slow process of convergence. After 20 rounds of fighting, Their beliefs still have not converged and the counterinsurgent still has no incentive to change strategies. Table 2 shows a similar pattern when certainty about means is low but certainty about strategy is high.[9]

But table 3 shows a different pattern entirely. Here, the counterinsurgent is more certain about means than she is about strategy.[10] Once again, there are a few periods during which expectations diverge before they begin to converge. But notably, they begin to converge again sooner, and the rate of convergence is much higher. Even more promisingly, 1 has the opportunity to recognize her error in choosing the better strategy and has an incentive to

[8]In this numerical example, certainty about both is low. Specifically, $\alpha_1^W = 1$, $\beta_1^W = 4$, $\alpha_1^{CC} = 3$, $\beta_1^{CC} = 1$, $\alpha_1^{CG} = 3$, $\beta_1^{CG} = 2$, $\alpha_1^{GG} = 2$, $\beta_1^{GG} = 1$, $\alpha_1^{GC} = 4$, $\beta_1^{GC} = 3$. The result is roughly the same when certainty about both is high.

[9]Specifically, $\alpha_1^W = 10$, $\beta_1^W = 40$, $\alpha_1^{CC} = 3$, $\beta_1^{CC} = 1$, $\alpha_1^{CG} = 3$, $\beta_1^{CG} = 2$, $\alpha_1^{GG} = 2$, $\beta_1^{GG} = 1$, $\alpha_1^{GC} = 4$, $\beta_1^{GC} = 3$.

[10]Specifically, $\alpha_1^W = 1$, $\beta_1^W = 4$, $\alpha_1^{CC} = 30$, $\beta_1^{CC} = 10$, $\alpha_1^{CG} = 30$, $\beta_1^{CG} = 20$, $\alpha_1^{GG} = 20$, $\beta_1^{GG} = 10$, $\alpha_1^{GC} = 40$, $\beta_1^{GC} = 30$.

t	$L_1^C(X_t^C, X_t^G, X_t^W, t)$	$L_1^G(X_t^C, X_t^G, X_t^W, t)$	$L_2^C(X_t^C, X_t^G, X_t^W, t)$	$L_2^G(X_t^C, X_t^G, X_t^W, t)$	$\|L_1^C(.) - L_2^C(.)\|$	$\|L_1^G(.) - L_2^G(.)\|$
0	0.837009	0.674445	0.275491	0.269134	0.561518	0.405311
1	0.791053	0.682159	0.208319	0.187652	0.582734	0.494507
2	0.699693	0.622449	0.164917	0.144621	0.534777	0.477828
3	0.603707	0.548195	0.136396	0.120721	0.467312	0.427474
4	0.519206	0.480075	0.116897	0.106058	0.402309	0.374017
5	0.450269	0.423555	0.102899	0.0958784	0.34737	0.327676
6	0.395848	0.37822	0.0923457	0.0879952	0.303502	0.290225
7	0.353242	0.341975	0.0840383	0.0814291	0.269203	0.260546
8	0.319648	0.312652	0.0772658	0.0757407	0.242383	0.236911
9	0.292726	0.28849	0.0715937	0.070721	0.221132	0.217769
10	0.270695	0.268189	0.0667455	0.0662551	0.203949	0.201934
11	0.252272	0.25082	0.0625373	0.0622662	0.189735	0.188554
12	0.236554	0.235729	0.058841	0.0586933	0.177713	0.177036
13	0.22291	0.22245	0.0555636	0.0554844	0.167347	0.166965
14	0.210896	0.210643	0.0526353	0.0525934	0.158261	0.15805
15	0.200196	0.200059	0.0500017	0.0499798	0.150194	0.15008
16	0.190578	0.190506	0.0476198	0.0476086	0.142958	0.142897
17	0.18187	0.181833	0.0454549	0.0454492	0.136415	0.136384
18	0.173939	0.17392	0.0434784	0.0434756	0.130461	0.130444
19	0.166679	0.16667	0.0416667	0.0416654	0.125013	0.125005
20	0.160006	0.160002	0.04	0.0399994	0.120006	0.120002

Table 5.1: Equally certain about means and about strategy

t	$L_1^C(X_t^C, X_t^G, X_t^W, t)$	$L_1^G(X_t^C, X_t^G, X_t^W, t)$	$L_2^C(X_t^C, X_t^G, X_t^W, t)$	$L_2^G(X_t^C, X_t^G, X_t^W, t)$	$\|L_1^C(.) - L_2^C(.)\|$	$\|L_1^G(.) - L_2^G(.)\|$
0	0.837009	0.674445	0.275491	0.269134	0.561518	0.405311
1	0.842741	0.714565	0.208319	0.187652	0.634421	0.526913
2	0.823968	0.730592	0.164917	0.144621	0.659051	0.585971
3	0.797758	0.733411	0.136396	0.120721	0.661362	0.61269
4	0.771577	0.729061	0.116897	0.106058	0.65468	0.623003
5	0.748081	0.720912	0.102899	0.0958784	0.645182	0.625034
6	0.727735	0.710847	0.0923457	0.0879952	0.63539	0.622852
7	0.710164	0.699909	0.0840383	0.0814291	0.626125	0.61848
8	0.694773	0.688672	0.0772658	0.0757407	0.617507	0.612932
9	0.68101	0.677447	0.0715937	0.070721	0.609416	0.606726
10	0.668441	0.666395	0.0667455	0.0662551	0.601696	0.60014
11	0.656754	0.655596	0.0625373	0.0622662	0.594217	0.59333
12	0.645733	0.645089	0.058841	0.0586933	0.586892	0.586395
13	0.635238	0.634884	0.0555636	0.0554844	0.579674	0.579399
14	0.625173	0.624981	0.0526353	0.0525934	0.572538	0.572388
15	0.615477	0.615375	0.0500017	0.0499798	0.565476	0.565395
16	0.60611	0.606056	0.0476198	0.0476086	0.55849	0.558447
17	0.59704	0.597013	0.0454549	0.0454492	0.551585	0.551563
18	0.588248	0.588234	0.0434784	0.0434756	0.54477	0.544759
19	0.579716	0.57971	0.0416667	0.0416654	0.53805	0.538044
20	0.571432	0.571428	0.04	0.0399994	0.531432	0.531429

Table 5.2: Uncertain about means and certain about strategy

t	$L_1^C(X_t^C, X_t^G, X_t^W, t)$	$L_1^G(X_t^C, X_t^G, X_t^W, t)$	$L_2^C(X_t^C, X_t^G, X_t^W, t)$	$L_2^G(X_t^C, X_t^G, X_t^W, t)$	$\lvert L_1^C(.) - L_2^C(.)\rvert$	$\lvert L_1^G(.) - L_2^G(.)\rvert$
0	0.982904	0.893949	0.275491	0.269134	0.707413	0.624815
1	0.962673	0.914695	0.208319	0.187652	0.754354	0.727043
2	0.937618	0.918283	0.164917	0.144621	0.772701	0.773662
3	0.907463	0.908055	0.136396	0.120721	0.771068	0.787334
4	0.87222	0.885872	0.116897	0.106058	0.755323	0.779813
5	0.832182	0.853173	0.102899	0.0958784	0.729283	0.757294
6	0.787898	0.811445	0.0923457	0.0879952	0.695552	0.72345
7	0.740136	0.76239	0.0840383	0.0814291	0.656097	0.68096
8	0.689826	0.707922	0.0772658	0.0757407	0.61256	0.632182
9	0.638	0.650084	0.0715937	0.070721	0.566407	0.579363
10	0.585722	0.590907	0.0667455	0.0662551	0.518976	0.524652
11	0.534024	0.532285	0.0625373	0.0622662	0.471487	0.470019
12	0.483854	0.475856	0.058841	0.0586933	0.425013	0.417162
13	0.436031	0.422924	0.0555636	0.0554844	0.380467	0.36744
14	0.391214	0.374431	0.0526353	0.0525934	0.338578	0.321838
15	0.34989	0.33095	0.0500017	0.0499798	0.299888	0.28097
16	0.312371	0.292721	0.0476198	0.0476086	0.264751	0.245112
17	0.278805	0.259704	0.0454549	0.0454492	0.23335	0.214255
18	0.249192	0.231642	0.0434784	0.0434756	0.205714	0.188166
19	0.223409	0.208122	0.0416667	0.0416654	0.181742	0.166456
20	0.201237	0.188641	0.04	0.0399994	0.161237	0.148641

Table 5.3: Certain about means and uncertain about strategy

switch strategies after 4 periods. In fact, from period 5 onward the sequence may no longer be valid — by switching tactics, she may no longer end up fighting ambiguous battles, and when she does, she will win according to G, not C.

The divergence in expectations that occurs in all three cases is noteworthy. Kristopher Ramsay has suggested that the empirical evidence stands in direct contradiction to the bargaining literature that models war as a means of learning overcoming mutual optimism. If this were true, he infers, then the probability of reaching a settlement should go up with each battle fought. But in his large-N study he finds that in general this probability initially goes down, before eventually rebounding. The model above is perhaps a means of reconciling Ramsay's finding with the bargaining literature.

5.4 Conclusion

The goal of this chapter has been to construct a plausible model of the relationship between domestic politics and conflict during an insurgency and counterinsurgency. As chapter 2 suggested, there is little evidence of a relationship between *US* domestic politics and the insurgency and counterinsurgency in Iraq — except, perhaps, insofar as the expectation of an inexorable decline in US public opinion is a necessary precondition for the very existence of the insurgency. Otherwise, there is no evidence that Iraqi insurgents pay any attention whatsoever to fluctuations in US support for the counterinsurgency. And while there is some evidence that US counterinsurgents did pay some attention to these fluctuations, it may well be that they shouldn't have, as they don't appear to have been able to have any influence upon it.

But what about the support of domestic political actors in Iraq? Theorists of insurgency and counterinsurgency alike seem to agree that the importance of the support of the civilian population is a key factor that differentiates insurgency and counterinsurgency from other kinds of conflict. This potential for support, besides the flexibility irregular armies, is widely viewed as the primary (some would say only) advantage that an insurgent enjoys. Yet there is disagreement over what, if anything, the counterinsurgent can do to counteract this advantage.

This is, ultimately, a disagreement over how to measure a war's success. And given how difficult it is to directly measure the support of Iraqi civilians, let alone determine the extent to which such support influences the belliger-

ents' ultimate probabilities of success, this disagreement cannot be quickly and easily decided. If a counterinsurgent can and must win over the civilian population to its side in order to prevail, but does not realize that this is the case, then it can lose the war even when apparently winning decisive victories on the battlefield. The model above is an attempt to sketch some of the implications of this approach to thinking about insurgency and counterinsurgency. In particular, it suggests that a counterinsurgent can only learn and adapt during war if she is relatively less confident about her strategy than she is about the means at her disposal. It also suggests that, at least initially, each battle can make the belligerents less rather than more willing to settle.

This kind of disagreement can help to explain the increasing violence in Iraq beginning with the US invasion and ending with the "surge." During this period of time, US and insurgent leadership alike did, in fact, make mutually optimistic statements not only about future but past successes. Thus it was possible in 2005 for Sheikh Harth Al-Dhari, the head of the Sunni Clerics Association in Iraq and prominent advocate of the insurgency, to speak of the "failure of [America's] enterprise in Iraq" while US Vice President Dick Cheney suggested that the insurgency was in its "last throes" and US Major General Ray Odierno claimed the insurgents had been "brought to their knees" and that "within six months you're going to see some normalcy" (Al-Dhari, 2005; Sanger, 2006; Ricks, 2006). Similarly, in 2006, US President George W. Bush claimed that "Terrorists are losing on the field of battle, so they are fighting this war through the pictures we see on television and in the newspapers

every day. They're hoping to shake our resolve and force us to retreat. They are not going to succeed," while self-proclaimed head of "Al Qaeda in Iraq" Abu Mus'ab Al-Zarqawi made similar claims about Bush (Al-Zarqawi, 2006; Sanger, 2006):

> You have become an utter liar. You deceive your people and supporters. Whenever the mujahidun increase their blows, you increase your lies and falsehoods, claiming that the situation is under control. But then it becomes clear to all that what you say is a lie, and so on. You are like someone who tries to cure his drunkenness with alcohol. You have not been honest with yourself or your people for a single moment, even though your forefathers had an integrity you lack. Why don't you reveal the truth about your soldiers, and the collapse of their motivation to fight, so that your people will know the truth about the war? Why don't you tell them about the continuing suicides among your soldiers? Why don't you tell them that none of your soldiers go to sleep unless they take sleeping pills and hallucinatory drugs, which make them lose their minds, and they become like dumb beasts, led by the pro-Zionist and Evangelist generals of your war to the slaughterhouses of perdition? Why don't you tell them about your soldiers' collective fleeing, and about the growing rebelliousness among their ranks?

That the kinds of disagreement I have attempted to model in this chapter might make belligerents less willing to settle may have some bearing on

recent findings in the empirical literature on conflict in general. Ramsay, as mentioned above, has argued that the evidence suggests that *in general* the probability of a settlement between belligerents goes up rather than down with each initial battle fought, though this probability ultimately peaks and then declines with subsequent battles. This pattern can perhaps be explained by identifying some analogue in conventional warfare to the kind of disagreement I have described here, over the proper metrics for success. In the conclusion I will discuss this and other possible applications and extensions of this model as well as some empirical tests that the model suggests.

Chapter 6

Conclusion

The current project is an attempt to explain how domestic politics might affect disagreements between potential belligerents in a conflict setting when such disagreement is due to differences in *beliefs* rather than *information*. I have endeavored to explain and show evidence of many of the shortcomings of models of domestic politics' influence on international conflict that treat said influence as purely informational. I have then developed a formal model that allows for domestic politics to affect conflict outcomes not by communicating otherwise private information, but by allowing for belligerents to develop different beliefs about public information. I summarize each of these contributions below.

6.1 Summary of findings

6.1.1 No emboldenment effect in Iraq

The first contribution of the current project is to suggest that, contrary to the assertions of the Bush administration and its supporters, and to many of the implications deriving from the "audience costs" literature, there is little evidence to support the argument that domestic politics in the US — in

particular US public opinion — have had any direct, immediate effect on the dynamics of the war in Iraq.

The one notable exception is that US forces seem to be responsive to fluctuations in public opinion. This would appear to be evidence that the Bush administration was using US foreign policy as an instrument to influence domestic electoral outcomes. However, the evidence also seems to suggest that the Bush administration was largely unsuccessful in this regard. This caveat aside, there seems to be little in the way of a relationship between US domestic politics and the war in Iraq.

This disconnect can be explained as follows. First, there are good reasons for political elites in both the US and Iraq to ignore the short–run fluctuations in American public opinion of the war, as these fluctuations have little to do with the dynamics of the war itself. There is no observable correlation between the two — fluctuations in American public opinion of the war appear to be more or less independent of the realities on the ground in Iraq. Thus, because public opinion is uninformed, it is — in the short–run, at least — essentially uninformative.

Second, disagreements among belligerents need not be exclusively informational. When disagreement is due also to differences in beliefs about how to interpret available information, the causal mechanisms identified by the "audience costs" literature can be confounded. Thus, while there may have been some useful information for Iraqi insurgents to glean from an observation of US domestic politics — especially if, as suggested above, US forces were respon-

sive to US domestic audiences — this potential for learning was undermined by belligerents' different beliefs about how to interpret that information.

6.1.2 Disagreements about success and failure

This leads to the current project's second and more important contribution. It is my contention that treating disagreements leading to conflict as having less to do with asymmetries in available information and more to do with differences in beliefs about how to interpret available information implies a very different role for domestic politics in conflict settings. To do so opens up the possibility of domestic audiences contributing not only to overcoming disagreement and thus avoiding conflict or hastening its termination, but also to increasing disagreement, and thus increasing the likelihood and duration of war. Given different metrics of success, belligerents might reasonably disagree about battlefield outcomes — i.e. about who has won and who has lost. Under such circumstances, fighting increases rather than decreases mutual optimism.

6.2 Implications

If the relationship — or lack thereof — observed between US domestic politics and belligerents in Iraq is observed in other conflict settings as well, this would have some substantial bearing on much of the current literature on bargaining and war. In particular, the now more or less canonical "audience costs" models developed by Kenneth Schultz and the "selectorate" models Bruce Bueno de Mesquita and his coauthors rely for their results on precisely

the kind of short–run dynamics that my empirical analysis suggests matter very little in the Iraq war. When it is the case that only long–run trends in domestic politics have any direct bearing on conflict dynamics, then the "audience costs" model of international politics overstates the informational contribution to conflict bargaining settings of domestic politics. A foreign adversary need only note, for example, that US political institutions allow the voting public to remove from office foreign policymakers whose decisions displease them and that, as Saddam Hussein has suggested and John Mueller has evinced, the same voting public has little stomach for protracted war, and she will have a very good notion of the future direction of US foreign policy. She need not actually pay any attention what is actually happening domestically, because neither of these observations requires much in the way of additional confirmatory evidence, nor does either imply that there is anything additional to learn from following domestic politics.

That empirical research such as Iyengar and Monten appears to find evidence of the very relationship that the current study suggests does not obtain is instructive. In the introduction and in chapter 2, I stressed the importance in analyzing conflict settings of accounting for the possibility of simultaneity or multiple endogeneity — i.e., that any given variable may be at once as cause and an effect. Here, it turns out, is a case in point. As I have shown, US forces are responsive to fluctuations in American public opinion of the war, and Iraqi insurgents are responsive to fluctuations in the behavior of US forces. The appearance of an "emboldenment effect" in Iraq is thus

likely an artifact of a failure both to account for US forces' behavior. That is, what some see as insurgents' responsiveness to public opinion could instead be merely insurgents' responsiveness to US forces who are themselves responsive to US domestic politics. This, then, is a powerful argument for taking seriously the possibility of multiple endogeneity in conflict settings.

Finally, and most importantly, it is my contention that the role of domestic politics in international conflict settings is not always as positive as the "audience costs" literature suggests. When access to public information about domestic audiences is limited, belligerents can reasonably disagree about the nature of their influence on conflict outcomes. This can cause them to become more rather than less mutually optimistic about their likelihoods of winning, thus protracting conflict. My formal model of this possibility suggests that under such circumstances, overcoming mutual optimism and thus ending conflict is possible (in fact, given sufficient time, it is inevitable), but much harder to do.

6.3 Future research

Three avenues for future scholarship immediately suggest themselves.

First, the full implications of the formal model developed in the previous chapter cannot be determined by analyzing a single numerical example, as I have done in the previous chapter. The next logical step is to come up with some analog to comparative statics in standard game theoretical analysis. This will involve coming up with some generalized measure of the rate of

convergence of belligerents' expectations, and then examining how differences in initial beliefs influence this rate.

Second, the relationships suggested by the formal model can and should be tested empirically. While the data I have on the Iraq war might be suitable for this purpose, there are of course serious problems in using the same data for model building and model testing [Leamer]. Looking at another war is the next logical step.

Third, and relatedly, the relationship between US public opinion and the behavior of US forces in Iraq is interesting, and bears further analysis. My discussion of this result suggested that it might be explained in part by the influence of academics on foreign policy — what might be called the "Peter Feaver Effect." I have presented some evidence for such an effect in chapter 4, but this evidence is inconclusive. One additional possible test of this effect might be to look at other conflicts involving the US to see if there is evidence of similar behavior where Feaver did not exert such direct influence over policy.[1]

Fourth and finally, while I have limited the scope of the current project to the explanation of insurgency and counterinsurgency, the formal model I have developed may well have other applications to war in general. I have suggested above that domestic audiences might be directly or indirectly responsible for disagreements among belligerents about how best to measure success and failure. This requires that there be room for disagreement about

[1]The idea of looking for evidence of a "Peter Feaver Effect" is due to Patrick McDonald (private coversation).

the nature of domestic audiences' influence on conflict outcomes. What the nature of such disagreement might be outside of the context of insurgency and counterinsurgency is an open question, and worth further inquiry.

Bibliography

Achen, C. (2002). Toward a New Political Methodology: Microfoundations and ART. *Annual Review of Political Science*, 5:423–450.

Al-Dhari, S. H. (2005). Sheikh Hareth Al-Dhari, Head of Sunni Clerics Association in Iraq: I Support the Resistance; Terrorism in Iraq Carried Out by U.S., Israel, and Government Militias. September 2. *MEMRI*.

Al-Zarqawi, A. M. (2006). New Video by Al-Qaeda Commander in Iraq Abu Mus'ab Al-Zarqawi. April 26. *MEMRI*.

Almond, G. (1965). *The American People and Foreign Policy*. Praeger Publishing, New York, NY.

Ashcroft, J. (2001). Testimony before the Senate Committee on the Judiciary. December 6.

Axelrod, R. (1981). The Emergence of Cooperation among Egoists. *American Political Science Review*, 75(2):306–318.

Baker, P. and Balz, D. (2005). Bush Words Reflect Public Opinion Strategy. June 30. *Washington Post*.

Baum, M. and Groeling, T. (2009). *War Stories*. Princeton University Press, Princeton, NJ.

Brandt, P. and Colaresi, M. (2008). The Dynamics of Reciprocity, Accountability, and Credibility. *Journal of Conflict Resolution*, 5(2):343–374.

Bueno de Mesquita, B. and Siverson, R. M. (1995). War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability. *American Political Science Review*, 89(4):841–855.

Bueno de Mesquita, B., Smith, A., Siverson, R. M., and Morrow, J. D. (2003). *The Logic of Political Survival*. The MIT Press, Cambridge, MA.

Bush, G. W. (2001). Address to a Joint Session of Congress and the American People. September 20.

Butler, J. (2002). Explanation and Exoneration, or What We Can Hear. *Social text*, 20(3):177–188.

Callwell, C. (1996). *Small Wars*. University of Nebraska Press, Lincoln, NE.

Chinh, T. (2001). *Resistence Will Win*. University of the Pacific Press.

Diamond, L. (2005). *Squandered Victory: The American Occupation and the Bungled Effort to Bring Democracy to Iraq*. Owl Books, New York, NY.

Downs, G. and Rocke, D. (1994). Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War. *American Journal of Political Science*, 38(2):362–380.

Elster, J. (2007). *Explaining Social Behavior*. Cambridge University Press, Cambridge, UK.

Fearon, J. (1994). Domestic Political Audiences and the Escalation of International Disputes. *American Political Science Review*, 88(3):577–592.

Fearon, J. (1995). Rationalist Explanations for War. *International Organization*, 49(3):379–414.

Feaver, P. D. and Gelpi, C. (2005). *Choosing Your Battles: American Civil-Military Relations and the Use of Force*. Princeton University Press, Princeton, NJ.

Fey, M. and Ramsay, K. (2006). The Common Priors Assumption: A Comment on" Bargaining and the Nature of War". *Journal of Conflict Resolution*, 50(4):607–613.

Fey, M. and Ramsay, K. W. (2007). Mutual optimism and war. *American Journal of Political Science*, 51(4):738–754.

Freeman, J. R., Williams, J. T., and Lin, T.-m. (1989). Vector Autoregression and the Study of Politics. *American Journal of Political Science*, 33(4):842–877.

Galula, D. (2006). *Counterinsurgency Warfare*. Praeger Publishing, Toronto, Ontario.

Gartner, S. S. (1998). Opening Up the Black Box of War. *Journal of Conflict Resolution*, 42(3):252–258.

Gelpi, C., Reifler, J., and Feaver, P. (2007). Iraq the Vote: Retrospective and Prospective Foreign Policy Judgments on Candidate Choice and Casualty Tolerance. *Political Behavior*, 29(2):151–174.

Goemans, H. (2000). *War and Punishment: The Causes of War Termination and the First World War*. Princeton University Press, Princeton, NJ.

Goff, B. and Tollison, R. D. (1987). The Allocation of Death in the Vietnam War: A Public Choice Perspective. *Southern Economic Journal*, 54(2):316–321.

Goldstein, J. S. and Freeman, J. R. (1990). *Three-Way Street*. The University of Chicago Press, Chicago, IL.

Granato, J. and Scioli, F. (2004). Puzzles, Proverbs, and Omega Matrices: The Scientific and Social Significance of Empirical Implications of Theoretical Models (EITM). *Perspectives on Politics*, 2(2):313–323.

Granger, C. (1969). Invesitigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37(3):424–438.

Guisinger, A. and Smith, A. (2002). Honest Threats: The Interaction of Reputation and Political Institutions in International Crises. *Journal of Conflict Resolution*, 46:175–200.

Gujarati, D. (2002). *Basic Econometrics*. McGraw-Hill, Colombus, OH.

Harsanyi, J. (1967). Games with Incomplete Information Played by Bayesian Players, Part I. *Management Science*, 14(3):159–182.

Harsanyi, J. (1968a). Games with Incomplete Information Played by Bayesian Players, Part II. *Management Science*, 14(5):320–334.

Harsanyi, J. (1968b). Games with Incomplete Information Played by Bayesian Players, Part III. *Management Science*, 14(7):486–502.

Hashim, A. S. (2006). *Insurgency and Counter–Insurgency in Iraq*. Cornell University Press, Ithaca, NY.

Horowitz, D. (2003). The Fifth Column Left Declares War. *Front Page Magazine*.

Iyengar, R. and Monten, J. (2008). Is There an "Emboldenment" Effect? Evidence from the Insurgency in Iraq. *NBER Working Paper*.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2):263–292.

Keefer, B. (2002a). Quieting the Homefront. March 3. *Slate*.

Keefer, B. (2002b). The Patriot Patrol Strikes Again! August 6. *Slate*.

Kelly, M. (2001, September 26). ... Pacifist claptrap. September 26. *Washington Post*.

Kilcullen, D. J. (2009). *The Accidental Guerrilla*. Oxford University Press, Oxford, UK.

Kupchan, C. and Trubowitz, P. (2007). Dead Center: The Demise of Liberal Internationalism in the United States. *International Security*, 32(2):7–44.

Kydd, A. H. (2005). *Trust and mistrust in international relations*. Princeton University Press, Princeton, NJ.

Lake, D. and Powell, R. (1999). *Strategic choice and international relations*. Princeton University Press, Princeton, NJ.

Larson, E. V. (1996). *Casualties and Consensus: The Historical Role of Casualties in Domestic Support for U.S. Military Operations*. RAND, Santa Monica, CA.

Lippman, W. (1922). *Public Opinion*. Free Press, New York, NY.

Machiavelli, N. (2005). *The Prince*. Bedford Saint Martin's Press, Boston, MA.

Mackesy, P. (1993). *The War for America*. Bison Books, Lincoln, NE.

Mearsheimer, J. (2011). *Why Leaders Lie: The Truth About Lying in International Politics*. Oxford University Press, Oxford, UK.

Mearsheimer, J. and Walt, S. (2006). The Israel Lobby. *London Review of Books*, 28(6).

Metz, S. (2008). *Iraq and the Evolution of American Strategy*. Potomac Books, Dulles, VA.

Morgenthau, H. (1967). *Politics Among Nations*. Knopf, New York, NY.

Morris, S. (1995). The Common Prior Assumption in Economic Theory. *Economics and Philosophy*, 11:227–227.

Mueller, J. (1973). *War, Presidents and Public Opinion*. John Wiley and Sons, Inc., New York, NY.

Mueller, J. (1994). *Policy and opinion in the Gulf War*. The University of Chicago Press, Chicago, IL.

Nagl, J. A. (2005). *Learning to Eat Soup with a Knife: Counterinsurgency Lessons from Malaysia and Vietnam*. The University of Chicago Press, Chicago, IL.

Nagl, J. A. (2007). Forward. In *The U.S. Army/Marine Corps Counterinsurgency Field Manual*. The University of Chicago Press, Chicago, IL.

Nye, J. (2009). Scholars on the Sidelines. April 13. *Washington Post*.

Nyhan, B. (2002). Bully brigade. March 4. *Slate*.

Olson, M. (1971). *The logic of collective action*. Harvard University Press, Cambridge, MA.

Orwell, G. (1968). London letter to Paris Review dated 12 July 1942. In Orwell, S. and Angus, I., editors, *Collected Essays, Journalism and Letters of George Orwell*. Harcourt, Brace, and World Inc., New York, NY.

Powell, R. (1991). Absolute and relative gains in international relations theory. *American Political Science Review*, 85(4):1303–1320.

Powell, R. (1994). Anarchy in international relations theory: the neorealist-neoliberal debate. *International Organization*, 48(2):313–344.

Powell, R. (2004). Bargaining and Learning While Fighting. *American Journal of Political Science*, 48(2):344–361.

Putnam, R., Evans, P. B., and Jacobson, H. K. (1993). *Double-Edged Diplomacy: International Bargaining and Domestic Politics*. University of California Press, Berkeley, CA.

Ramsay, K. (2008). Settling It on the Field: Battlefield Events and War Termination. *Journal of Conflict Resolution*, 52(6):850–879.

Reiter, D. (2011). *How Wars End*. Princeton University Press, Princeton, NJ.

Reiter, D. and Stam, A. C. (2002). *Democracies at War*. Princeton University Press, Princeton, NJ.

Richardson, L. F. (1960). *Arms and Insecurity*. Boxwood Press, Pacific Grove, CA.

Ricks, T. E. (2006). *Fiasco: The American military adventure in Iraq*. The Penguin Press.

Ricks, T. E. (2008). Then and Now. March 16. *Washington Post*.

Robinson, L. (2008). *Tell Me How this Ends: General David Petraeus and the Search for a Way Out of Iraq*. Public Affairs, New York, NY.

Roper Center (2007). Presidential Job Approval Ratings. In *Presidential Job Approval Ratings*.

Rosato, S. (2003). The flawed logic of democratic peace theory. *American Political Science Review*, 97(4):585–602.

Rose, G. (1998). Neoclassical Realism and Theories of Foreign Policy. *World Politics*, 51(1):144–172.

Sanger, D. (2006). Bush, Conceding Problems, Defends Iraq War. March 14. *New York Times*.

Sanger, D. (2007). Bush Adds Troops in Bid to Secure Iraq. January 10. *New York Times*.

Sanger, D. and Gordon, M. (2006). Options weighed for surge in G.I.'s to stabilize Iraq. December 16. *New York Times*.

Sartori, A. E. (2005). *Deterrence by Diplomacy*. Princeton University Press, Princeton, NJ.

Schelling, T. (1960). *The Strategy of Conflict*. Harvard University Press, Cambridge, MA.

Schornig, N. and Lembcke, A. C. (2006). The Vision of War without Casualties: On the Use of Casualty Aversion in Armament Advertisements. *Journal of Conflict Resolution*, 50(2):204–227.

Schrodt, P. A. (2006). TABARI Version 0.5. Technical report.

Schroeder, P. (1992). Did the Vienna Settlement Rest on a Balance of Power? *American Historical Review*, 97(3):683–706.

Schroeder, P. (1994). Historical Reality vs. Neo-Realist Theory. *International Security*, 19(1):104–148.

Schultz, K. (1998). Domestic Opposition and Signaling in International Crises. *American Political Science Review*, 92(4):829–844.

Schultz, K. (1999). Do Democratic Institutions Constrain or Inform?: Contrasting Two Institutional Perspectives on Democracy and War. *International Organization*, 52(2):233–266.

Schultz, K. (2001a). *Democracy and Coercive Diplomacy*. Cambridge University Press, Cambridge, UK.

Schultz, K. (2001b). Looking for Audience Costs. *Journal of Conflict Resolution*, 45(1):32–60.

Shane, S. (2005). Bush's Speech on Iraq War Echoes Voice of an Analyst. December 4. *New York Times*.

Sims, C. (1972). Money, Income and Causality. *American Economic Review*, 62(4):540–552.

Sims, C. (1980). Macroeconomics and Reality. *Econometrica*, 48(1):1–48.

Slantchev, B. (2011). Mutual Optimism as a Rationalist Explanation of War. *American Journal of Political Science*, 55(1):135–148.

Smith, A. (1998). International Crises and Domestic Politics. *American Political Science Review*, 92(3):623–638.

Smith, A. and Stam, A. (2004). Bargaining and the Nature of War. *Journal of Conflict Resolution*, 48(6):783–813.

Smith, T. (2009). Wilsonianism after Iraq. In Ikenberry, G. J. and Knock, T. J., editors, *The crisis of American foreign policy*. Princeton University Press, Princeton, NJ.

Stein, J. G. (1992). Deterrence and Compellence in the Gulf, 1990-91: A Failed or Impossible Task? *International Security*, 17(2):147–179.

Steort, L. (2006). Can Iran be Contained? October 23. *National Review*.

Stewart, J. (2010). The Daily Show. August 25. Television.

Sullivan, A. (2001). Why Did it Have to be a Perfect Morning? September 16. *Sunday Times of London*.

Trubowitz, P. (1998). *Defining the National Interest*. University of Chicago Press, Chicago, IL.

van Creveld, M. (2011). *The Age of Airpower*. Public Affairs, New York, NY.

von Clausewitz, C. (1993). *On War*. Everyman's Library, New York, NY.

Wagner, R. H. (1993). What was bipolarity? *International Organization*, 47(1):77–106.

Wagner, R. H. (2000). Bargaining and war. *American Journal of Political Science*, 44(3):469–484.

Wagner, R. H. (2007). *War and the State: The Theory of International Politics*. University of Michigan Press, Ann Arbor, MI.

Wagner, R. H. (2010). Rationalism and Security. *International Studies Encyclopedia*.

Walt, S. (1999). Rigor or Rigor Mortis?: Rational Choice and Security Studies. *International Security*, 23(4):5–48.

Walt, S. (2005). The Relationship between Theory and Policy in International Relations. *Annual Review of Political Science*, 8(1):23–48.

Waltz, K. (1979). *Theory of international politics*. Addison-Wesley, Boston, MA.

Willbanks, J. M. (2007). *The Tet Offensive: A Concise History*. Columbia University Press, New York, NY.

Wittman, D. (1979). How a War Ends: A Rational Model Approach. *Journal of Conflict Resolution*, 23(4):743–763.

Zeileis, A., Kleiber, C., Kramer, W., and Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics and Data Analysis*, 44(1):109–123.

Vita

Matthew Leonard Cohen was born in Orange, California on August 31, 1977, the son of Mary and Robert Cohen. His family moved soon thereafter to Santa Rosa, California, where he spent most of his childhood. In 1995, he left high school early to attend Bard College at Simon's Rock, the first and only four–year college dedicated to providing a liberal arts education to younger scholars. He graduated cum laude from Bard College at Simon's Rock in 1999, with a concentration in Politics, Law, and Society. In 2000, he joined the Peace Corps and served as a middle and high school teacher in Aktobe, Kazakhstan for two years, during which time he learned to speak the Russian language with intermediate fluency. Upon completing his service, he enrolled as a graduate student in the Department of Government at the University of Texas in Austin, where he won the prestigious University and MacDonald Fellowships at the University of Texas and the Clogg Fellowship at the Interuniversity Consortium for Political and Social Research Summer Program at the University of Michigan in Ann Arbor. While completing his dissertation, he was a teaching assistant in courses in American politics, international relations, and political methodology, and was appointed as an assistant instructor in 2007, teaching his own course, in American politics. In 2009, while still working on his dissertation, he was hired as a professor in the Social Sciences Department at Bard High School Early College in Queens, New York, a

joint venture between the New York City Department of Education and Bard College, where students complete an accelerated high school program in their first two years, and then spend their last two years enrolled in a rigorous liberal arts college curriculum.

Permanent address: 61 2nd Place, #2
Brooklyn, NY 11231

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.